

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 2, 2003, 12:17:07 ; Search time 4718 Seconds
(without alignments)
9517.938 Million cell updates/sec

Title: US-09-891-138A-1
Perfect score: 1543
Sequence: 1 gctcctggcagagttttctg.....tgccataataaataatata 1543

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pi.*
9: gb_pl.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pi.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
C	1	1543	100.0	1543	6	AX376573	Sequence
	2	1533.4	99.4	1598	10	AF295367	Mus muscu
	3	645.8	41.9	60298	2	AC116149	Mus muscu
	4	633.6	41.1	60298	2	AC116149	Mus muscu
	5	592.4	38.4	1325	9	AF247785	Homo sapi
	6	592.4	38.4	1380	9	AF348078	Homo sapi
	7	592.4	38.4	1449	9	BC030948	Homo sapi
	8	592.4	38.4	1542	6	AX342665	Sequence
	9	590.2	38.3	90343	9	AC116026	Homo sapi
	10	590.2	38.3	132745	9	AC068647	Homo sapi
c	11	589.2	38.2	1996	6	AF035943	Sequence
	12	588.6	38.1	158144	2	AC016455	Homo sapi
	13	147.6	9.6	1977	5	AF031897	Meleagris
	14	145.2	9.4	69462	2	AC101335	Mus muscu
	15	126.6	8.2	1014	6	AX148186	Sequence
	16	126.6	8.2	1014	6	AX379468	Sequence
	17	126.6	8.2	1014	6	AX384211	Sequence
	18	126.6	8.2	1014	9	AB083598	Homo sapi
	19	126.6	8.2	1014	9	AF411109	Homo sapi
	20	126.6	8.2	1081	6	AX458238	Sequence
c	21	126.6	8.2	1092	9	AF370886	Homo sapi
	22	126.6	8.2	1414	9	AB065877	Homo sapi
	23	126.6	8.2	1729	6	AX191332	Sequence
	24	126.6	8.2	9905	6	AX379470	Sequence
	25	126.6	8.2	67645	9	AL356486	Human DNA
	26	126.6	8.2	156555	9	AC026756	Homo sapi
	27	125	8.1	1014	6	AX05130	Sequence
	28	125	8.1	1014	6	AX464561	Sequence
	29	119.2	7.7	1020	6	AX147840	Sequence
	30	114	7.4	136930	5	AL590151	Sequence
c	31	113.6	7.4	1163	5	GDP2Y3	Zebrafish
	32	109	7.1	202838	2	AC108794	G. domestic
	33	108.4	7.0	91107	2	AC116265	Mus muscu
	34	107	6.9	1074	5	AF069555	Rattus no
	35	104	6.7	1020	6	AX458308	Meleagris
	36	104	6.7	1076	6	AX375230	Sequence
	37	103.2	6.7	1498	5	GDPATREC	Sequence
	38	102.6	6.6	3204	10	RNU22830	Rattus norv
	39	101.6	6.6	1308	4	BTU34041	Bos taurus
	40	101.6	6.6	1666	4	BTP2YRECP	B. taurus mr
	41	100.8	6.5	1312	9	S81950	P2 purinoc
	42	100.8	6.5	1426	9	HSU42029	P2Y1
	43	100.8	6.5	2424	9	HSU42030	Human P2Y1
	44	100.8	6.5	16178	9	HS4569	Human P2Y
	45	100.8	6.5	176170	2	AC021662	Homo sapi

ALIGNMENTS

RESULT 1	AX376573	AX376573	Sequence 1	1543 bp	DNA	linear	PAT 01-MAR-2002
LOCUS	AX376573	Sequence 1	from Patent WO0200719.				
DEFINITION	AX376573	Sequence 1	from Patent WO0200719.				
ACCESSION	AX376573	Sequence 1	from Patent WO0200719.				
VERSION	AX376573.1	GI:19170674					
KEYWORDS	house mouse.						
SOURCE	Mus musculus						
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
AUTHORS	Lin, D.C., Zhao, J., Chen, J.L. and Cutler, G.						
TITLE	Novel receptors						
JOURNAL	Patent: WO 0200719-A 1 03-JAN-2002;						

Pred. No. is the number of results predicted by chance to have a

FEATURES		Tularik Inc. (US)		Location/Qualifiers	
source		1. 1543		/organism="Mus musculus"	
				/db_xref="taxon:10090"	
CDS		44..997		/note="mouse TGR18 G-protein coupled receptor (GPCR)"	
				/codon_start=1	
				/db_xref="GI:19170675"	
				/translation="MAQNSLSCNWLATEAILNKYLSAFYAFIEFIFGLLGNVTVFGY	
				LFCMKNWNSNLYFNLSIDFAPLPLIKSYANDKGTGDIKLSNRYVHLN	
				YTSILFTFTSMORYLKMYPFHEHFLQKKEFALLISLAWMLVLEVLPLMFLINSV	
				PKEGSNCIDYASGNENPHNLISLCLTLGLPLISVMGFFYKMWFLKRRSQQQA	
				TALPLDPQLVLAIVFISILFTPYHIMRNLRIRASRLDSWPOGCTQKAISKIYILTR	
				PLAFNSAINPIFVFLMGDHYREMLISKPERQKSLTSPT"	
BASE COUNT		438 a		352 c 293 g 460 t	
ORIGIN		Query Match		100.0%; Score 1543; DB 6; Length 1543;	
		Best Local Similarity		100.0%; Pred. No. 0;	
		Matches 1543; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GCTCTGGCAGAGTTTCTGTCGACACAGAGCCGACAGCAGAGATGGCACAGAATTATC	60		
DB	1	GCTCTGGCAGAGTTTCTGTCGACACAGAGCCGACAGCAGAGATGGCACAGAATTATC	60		
QY	61	TTGTGAGAAATGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA	120		
DB	61	TTGTGAGAAATGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA	120		
QY	121	TGCAATCGAGTTCATTTTGGAGTCTGTTGGGAATGCTACTGTGGTGTGCGCTACCTCTT	180		
DB	121	TGCAATCGAGTTCATTTTGGAGTCTGTTGGGAATGCTACTGTGGTGTGCGCTACCTCTT	180		
QY	181	CTGCATGAGAACTGGAACAGCAGAAATGCTATCTTTTAACTCTCTGACTT	240		
DB	181	CTGCATGAGAACTGGAACAGCAGAAATGCTATCTTTTAACTCTCTGACTT	240		
QY	241	TGCTTTCTGTGACCCCTTCCCATCTGATAAAGAGTTATGCCAATGATAAGGGACCTA	300		
DB	241	TGCTTTCTGTGACCCCTTCCCATCTGATAAAGAGTTATGCCAATGATAAGGGACCTA	300		
QY	301	TGGAGATGTTCTGTATAGCAACCGATATGCTTACACCAACCTCTACACCGCAT	360		
DB	301	TGGAGATGTTCTGTATAGCAACCGATATGCTTACACCAACCTCTACACCGCAT	360		
QY	361	CTCTTCCCTCACTTTTCATTTAGCATGGACCGATATCTGCTCATGAAAGTACCCCTTTCCGAGA	420		
DB	361	CTCTTCCCTCACTTTTCATTTAGCATGGACCGATATCTGCTCATGAAAGTACCCCTTTCCGAGA	420		
QY	421	ACACTTTCTACAAAAGAGAAATTTGCCATTTTAAATCTGCGTGGCTGTCTGGGCCCTTAGT	480		
DB	421	ACACTTTCTACAAAAGAGAAATTTGCCATTTTAAATCTGCGTGGCTGTCTGGGCCCTTAGT	480		
QY	481	GACCTTTAGAGTTCTACCCATGCTCTTTCATCAATCTGCTGCCAAAGAGAGGGCAG	540		
DB	481	GACCTTTAGAGTTCTACCCATGCTCTTTCATCAATCTGCTGCCAAAGAGAGGGCAG	540		
QY	541	TAACTCATCGACTATGCAAGTCTGGAACCCCTGAACACAATCTCAITACAGCCCTCTG	600		
DB	541	TAACTCATCGACTATGCAAGTCTGGAACCCCTGAACACAATCTCAITACAGCCCTCTG	600		
QY	601	CTGACTTTTGTGGGCTTCTAATTCCTCTCTCTGTGATGCTTCTTCTACTACAAGAT	660		
DB	601	CTGACTTTTGTGGGCTTCTAATTCCTCTCTCTGTGATGCTTCTTCTACTACAAGAT	660		
QY	661	GGTAGTCTTTAAAGAGAGGAGCCAGCAGCAATGCTGCCCTGCCACTGGACAAACC	720		
DB	661	GGTAGTCTTTAAAGAGAGGAGCCAGCAGCAATGCTGCCCTGCCACTGGACAAACC	720		
QY	721	CCAAACGCTGGTGGTCTGGCGGTTGTGATCTTCTATACCTCTTACACCCCTATCATAT	780		

RESULT 2	AF295367	1598 bp	mRNA	linear	ROD 06-APR-2001
LOCUS	AF295367				
DEFINITION	Mus musculus G-protein coupled receptor GPR91 mRNA, complete cds.				
ACCESSION	AF295367				
VERSION	AF295367.1	GI:12711490			
KEYWORDS	Mus musculus.				
SOURCE	Mus musculus				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
	1 (bases 1 to 1598)				
REFERENCE	Wittenberger, T., Schaller, H.C. and Heilebrand, S.				
AUTHORS	An expressed sequence tag (EST) data mining strategy succeeding in				
TITLE	the discovery of new G-protein coupled receptors				
JOURNAL	J. Mol. Biol. 307 (3), 799-813 (2001)				
MEDELINE	21172992				
PUBMED	11273702				

REFERENCE	2 (bases 1 to 1598)									
AUTHORS	Wittenberger,T., Schaller,C.H. and Hellebrand,S.									
TITLE	Direct Submission									
JOURNAL	Submitted (14-AUG-2000) ZMNH, Institut fur									
	Entwicklungsneurobiologie, Martinistr. 52, Hamburg 20246, Germany									
FEATURES	Location/Qualifiers									
source	1..1598									
	/organism="Mus musculus"									
	/strain="C57BL"									
	/db_xref="taxon:10090"									
CDS	74..1027									
	/note="orphan receptor"									
	/codon_start=1									
	/product="G-protein coupled receptor GPR91"									
	/protein_id="AAK01867.1"									
	/db_xref="GI:127111491"									
	/translation="MAQNLSCENLATEALINKYVLSAFVAIEFIFGLLCNVTVFVGY LFCMKWNSNVVLFNLISIDFAFLCTLPLIKSYANDKGTGYDVLICISNRYVLIHTNYS YTMLLTVLSMDRYLLMKYPFEHLFKQKFAILLISLAWLVLTLEVLPLMTFTINSV PREGNSCIDYASGNEPHNLYISGLTLLGFLIPUSVMCFYKMKVFLIKRSQQQA PLALDKPQRLVAVLYFISLTPYHMRNLRIASLDSWPQGCQTQAKIKSVITLTLR TAPLFLNSAINPIFYFLMGDHYREMLISKFRQYFKSLTSFT"									
BASE COUNT	465 a 358 c 303 g 472 t									
ORIGIN										
	Query Match 99.4%; Score 1533.4; DB 10; Length 1598;									
	Best Local Similarity 99.6%; Pred. No. 0;									
	Matches 1537; Conservative 0; Mismatches 6; Indels 0; Gaps 0;									
QY	1	GCTCCTGCGACAGT	TTCTGTGCGACAGACGAGCGGACGACAGAAATGCGACAGAAATTATC	60						
Db	31	GCTCCTGCGACAGT	TTCTGTGCGACAGACGAGCGGACGACAGAAATGCGACAGAAATTATC	90						
QY	61	TTGTGAGAA	TGGTGGCACACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA	120						
Db	91	TTGTGAGAA	TGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA	150						
QY	121	TGCAATCGAGT	TCATTTTGGACTGCTTGGAAATGTCACGTGGTGTGGCTACCTCTT	180						
Db	151	TGCAATCGAGT	TCATTTTGGACTGCTTGGAAATGTCACGTGGTGTGGCTACCTCTT	210						
QY	181	CTGCATGAAGA	CTGGCAACGACGACAAATGTCTATCTTTTAACTCTTCCATCTCTGACCT	240						
Db	211	CTGCATGAAGA	CTGGCAACGACGACAAATGTCTATCTTTTAACTCTTCCATCTCTGACCT	270						
QY	241	TGCTTTCTCTG	CGACCTTCCCATCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA	300						
Db	271	TGCTTTCTCTG	CGACCTTCCCATCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA	330						
QY	301	TGGAGATG	TGTTCTGTATAAGCAACCGATATGTGCTTCACCAACCTCTACACGAGCAT	360						
Db	331	TGGAGATG	TGTTCTGTATAAAGCAACCGATATGTGCTTCACCAACCTCTACACGAGCAT	390						
QY	361	CCCTCTCC	TCACTTTCATTAGCAGCGATATCTGCTCATGAAGTACCCCTTCCGAGA	420						
Db	391	GCCTTTC	CTCACTGTCATTAGCAGCGATATCTGCTCATGAAGTACCCCTTCCGAGA	450						
QY	421	ACACTTTCT	ACAAAAGAAGATTTGCCATTTAACTCGCTGGCTGTCTGGCCCTTAGT	480						
Db	451	ACACTTTCT	ACAAAAGAAGATTTGCCATTTAACTCGCTGGCTGTCTGGCCCTTAGT	510						
QY	481	GACCTTTAG	AAGTTCTACCTTCACTTTCATCAATTTCTGCCAAAAGAGGGCAG	540						
Db	511	GACCTTTAG	AAGTTCTACCTTCACTTTCATCAATTTCTGCCAAAAGAGGGCAG	570						
QY	541	TAACTGCAT	CGACTATGCAAGTTCTGGAAACCCCTGAACAAATCTCATTTACAGCCTCTG	600						
Db	571	TAACTGCAT	CGACTATGCAAGTTCTGGAAACCCCTGAACAAATCTCATTTACAGCCTCTG	630						
QY	601	CTGACTTTG	TGGGCTCCCTAATTCCTCTCTGTGATGCTCTTCTACTACAAGT	660						
Db	631	CTGACTTTG	TGGGCTCCCTAATTCCTCTCTGTGATGCTCTTCTACTACAAGT	690						

Thu Apr 3 10:42:25 2003

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP24-540E9
2 (bases 1 to 60298)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,I.,
Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zaitoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE
JOURNAL
COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L24912
Center clone name: 540_E_9

NOTE: This record contains 77 individual
sequencing reads that have not been assembled into
contigs. Runs of N are used to separate the reads
and the order in which they appear is completely
arbitrary. Low-pass sequence sampling is useful for
identifying clones that may be gene-rich and allows
overlap relationships among clones to be deduced.
However, it should not be assumed that this clone
will be sequenced to completion. In the event that
the record is updated, the accession number will
be preserved.

657 756: gap of 100 bp
757 1426: contig of 670 bp in length
1427 1526: gap of 100 bp
1527 2210: contig of 684 bp in length
2211 2310: gap of 100 bp
2311 2997: contig of 687 bp in length
2998 3097: gap of 100 bp
3098 3786: contig of 689 bp in length
3787 3886: gap of 100 bp
3887 4577: contig of 691 bp in length
4578 4677: gap of 100 bp
4678 5357: contig of 680 bp in length
5358 5457: gap of 100 bp
5458 6150: contig of 693 bp in length
6151 6250: gap of 100 bp
6251 6817: contig of 567 bp in length
6818 6917: gap of 100 bp
6918 7615: contig of 698 bp in length
7616 7715: gap of 100 bp
7716 8412: contig of 697 bp in length

8413 3512: gap of 100 bp
8513 9198: contig of 686 bp in length
9199 9298: gap of 100 bp
9299 9988: contig of 690 bp in length
9989 10088: gap of 100 bp
10089 10768: contig of 680 bp in length
10769 10868: gap of 100 bp
10869 11524: contig of 656 bp in length
11525 11624: gap of 100 bp
11625 12242: contig of 618 bp in length
12243 12342: gap of 100 bp
12344 13040: contig of 698 bp in length
13041 13140: gap of 100 bp
13141 13829: contig of 689 bp in length
13830 13929: gap of 100 bp
13930 14647: contig of 718 bp in length
14648 14747: gap of 100 bp
14748 15451: contig of 704 bp in length
15452 15551: gap of 100 bp
15552 16247: contig of 696 bp in length
16248 16347: gap of 100 bp
16348 17028: contig of 681 bp in length
17029 17128: gap of 100 bp
17129 17802: contig of 674 bp in length
17803 17902: gap of 100 bp
17903 18593: contig of 691 bp in length
18594 18693: gap of 100 bp
18694 19375: contig of 682 bp in length
19376 19475: gap of 100 bp
19476 20082: contig of 607 bp in length
20083 20182: gap of 100 bp
20183 20875: contig of 693 bp in length
20876 20975: gap of 100 bp
20976 21650: contig of 675 bp in length
21651 21750: gap of 100 bp
21751 22427: contig of 677 bp in length
22428 22527: gap of 100 bp
22528 23238: contig of 711 bp in length
23239 23338: gap of 100 bp
23339 24028: contig of 690 bp in length
24029 24128: gap of 100 bp
24129 24803: contig of 675 bp in length
24804 24903: gap of 100 bp
24904 25603: contig of 700 bp in length
25604 25703: gap of 100 bp
25704 26357: contig of 654 bp in length
26358 26457: gap of 100 bp
26458 27140: contig of 683 bp in length
27141 27240: gap of 100 bp
27241 27946: contig of 706 bp in length
27947 28046: gap of 100 bp
28047 28734: contig of 688 bp in length
28735 28834: gap of 100 bp
28835 29536: contig of 702 bp in length
29537 29636: gap of 100 bp
29637 30324: contig of 688 bp in length
30325 30424: gap of 100 bp
30425 31130: contig of 706 bp in length
31131 31230: gap of 100 bp
31231 31910: contig of 680 bp in length
31911 32010: gap of 100 bp
32011 32691: contig of 681 bp in length
32692 32791: gap of 100 bp
32792 33482: contig of 691 bp in length
33483 33582: gap of 100 bp
33583 34274: contig of 692 bp in length
34275 34374: gap of 100 bp
34375 35081: contig of 707 bp in length
35082 35181: gap of 100 bp
35182 35861: contig of 680 bp in length
35862 35961: gap of 100 bp
35962 36660: contig of 699 bp in length
36661 36760: gap of 100 bp

[illegible]

QY	411	CTTTCCGAGAACACTTTCTACAAAAGAGAAATTTGCCATTTAAATCTCGCTGGCTGTCT	470
Db	4250	CTTTCCGAGAACACTTTCTACAAAAGAGAAATTTGCCATTTAAATCTCGCTGGCTGTCT	4309
QY	471	GGGCGCTTAGTGACCTTAGAAGTTCTACCATGCTCCTTCATCAATTTCTGTGCCAAAAG	530
Db	4310	GGGCGCTTAGTGACCTTAGAAGTTCTACCATGCTCCTTCATCAATTTCTGTGCCAAAAG	4369
QY	531	AAGAGGGCAGTAACTCGATCGACTATGCAAGTCTTGGAACCCCTGAACACAAATCTCATTT	590
Db	4370	AAGAGGGCAGTAACTCGATCGACTATGCAAGTCTTGGAACCCCTGAACACAAATCTCATTT	4429
QY	591	ACAGCCTCTGCGTGACTTTGTGGGCTTCCTAAATTCCTCTCTGTGATGCTCTTCT	650
Db	4430	ACAGCCTCTGCGTGACTTTGTGGGCTTCCTAAATTCCTCTCTGTGATGCTCTTCT	4489
QY	651	ACTACAAGATGGTAGTCTTTAAAGAGGAGGAGCGACGACGAACCTGCCTGCCAC	710
Db	4490	ACTACAAGATGGTAGTCTTTAAAGAGGAGGAGCGACGACGAACCTGCCTGCCAC	4549
QY	711	TGGACAAACCCCAAGCGCTGTGGTCTCTGGGGTGTGATCTCTCTATACTTTCACAC	770
Db	4550	TGGAC - AACCCAAAGCGCTGGGGTCTCTGNNNNNNNNNNNNNNNNNNNNNNNN	4608
QY	771	CTATCATATCATGCGCAATTGAGGATCGCCTCAGCGCTGGATGTGGCCACAGGAT	830
Db	4609	NN	4668
QY	831	GTACACAGAGGCCATCAAT 851	
Db	4669	NNNNNNNCGGAGATCTGAT 4689	
RESULT	4		
LOCUS	AC116149	60298 bp	DNA linear
DEFINITION	Mus musculus clone RP24-540E9, LOW-PASS SEQUENCE SAMPLING.		
ACCESSION	AC116149		
VERSION	AC116149.1	GI:19703273	
KEYWORDS	HTG; HTGS_PHASE0.		
SOURCE	Mus musculus.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	1. (bases 1 to 60298)		
	Birren, B., Linton, L., Nusbaum, C. and Lander, E.		

	Query Match	41.9%;	Score 645.8;	DB 2;	Length 60298;
	Best Local Similarity	84.0%;	Pred. No. 6.9e-135;		
	Matches 673;	Conservative 0;	Mismatches 127;	Indels 1;	Gaps 1;
Qy	51	AGAAATTTATCTGTGAGAAATTTGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCT	110		
Db	3890	AGATCTGATATCTCGCCCTGTGGTGGAAATCTCAGGCTATCTGGAATAAGTACTACCTCT	3949		
Qy	111	CTGCATTTTATGCAATCGAGTTTCATTTTGGACATGCTTGGGAATGTCACCTGTGGTGTTCG	170		
Db	3950	CTGCATTTTATGCAATCGAGTTTCATTTTGGACATGCTTGGGAATGTCACCTGTGGTGTTCG	4009		
Qy	171	GCTACCTCTTCTGCATGAAGAACCTGGAAACACGACCAATGCTATCTTTTAAACCTTTCCA	230		
Db	4010	GCTACCTCTTCTGCATGAAGAACCTGGAAACACGACCAATGCTATCTTTTAAACCTTTCCA	4069		
Qy	231	TCTCTGACTTTGCTTTTCTGTGCAACCTTCCATCTCCATATAACAGATATGCCAATGATA	290		
Db	4070	TCTCTGACTTTGCTTTTCTGTGCAACCTTCCATCTCCATATAACAGATATGCCAATGATA	4129		
Qy	291	AGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTCCTTACACCAACCTCT	350		
Db	4130	AGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTCCTTACACCAACCTCT	4189		
Qy	351	ACACCAGCATCCTCTTCCCTCACCTTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACC	410		
Db	4190	ACACCAGCATCCTCTTCCCTCACCTTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACC	4249		

TITLE
JOURNAL

Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome

COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L24912
 Center clone name: 540_E_9

 * NOTE: This record contains 77 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 656: contig of 656 bp in length
 657 756: gap of 100 bp
 757 1426: contig of 670 bp in length
 1427 1526: gap of 100 bp
 1527 2210: contig of 684 bp in length
 2211 2310: gap of 100 bp
 2311 2997: contig of 687 bp in length
 2998 3097: gap of 100 bp
 3098 3786: contig of 689 bp in length
 3787 3886: gap of 100 bp
 3887 4577: contig of 691 bp in length
 4578 4677: gap of 100 bp
 4678 5357: contig of 680 bp in length
 5358 5457: gap of 100 bp
 5458 6150: contig of 693 bp in length
 6151 6250: gap of 100 bp
 6251 6817: contig of 567 bp in length
 6818 6917: gap of 100 bp
 6918 7615: contig of 698 bp in length
 7616 7715: gap of 100 bp
 7716 8412: contig of 697 bp in length
 8413 8512: gap of 100 bp
 8513 9198: contig of 686 bp in length
 9199 9298: gap of 100 bp
 9299 9988: contig of 690 bp in length
 9989 10088: gap of 100 bp
 10089 10768: contig of 680 bp in length
 10769 10868: gap of 100 bp
 10869 11524: contig of 656 bp in length
 11525 11624: gap of 100 bp
 11625 12242: contig of 618 bp in length
 12243 12342: gap of 100 bp
 12343 13040: contig of 698 bp in length
 13041 13140: gap of 100 bp
 13141 13829: contig of 689 bp in length
 13830 13929: gap of 100 bp
 13930 14647: contig of 718 bp in length
 14648 14747: gap of 100 bp
 14748 15451: contig of 704 bp in length
 15452 15551: gap of 100 bp
 15552 16247: contig of 696 bp in length
 16248 16347: gap of 100 bp
 16348 17028: contig of 681 bp in length
 17029 17128: gap of 100 bp
 17129 17802: contig of 674 bp in length
 17803 17902: gap of 100 bp
 17903 18593: contig of 691 bp in length
 18594 18693: gap of 100 bp
 18694 19375: contig of 682 bp in length

19376 19475: gap of 100 bp
 19476 20082: contig of 607 bp in length
 20083 20182: gap of 100 bp
 20183 20875: contig of 693 bp in length
 20876 20975: gap of 100 bp
 20976 21650: contig of 675 bp in length
 21651 21750: gap of 100 bp
 21751 22427: contig of 677 bp in length
 22428 22527: gap of 100 bp
 22528 23238: contig of 711 bp in length
 23239 23338: gap of 100 bp
 23339 24028: contig of 690 bp in length
 24029 24128: gap of 100 bp
 24129 24803: contig of 675 bp in length
 24804 24903: gap of 100 bp
 24904 25603: contig of 700 bp in length
 25604 25703: gap of 100 bp
 25704 26357: contig of 654 bp in length
 26358 26457: gap of 100 bp
 26458 27140: contig of 683 bp in length
 27141 27240: gap of 100 bp
 27241 27946: contig of 706 bp in length
 27947 28046: gap of 100 bp
 28047 28734: contig of 688 bp in length
 28735 28834: gap of 100 bp
 28835 29536: contig of 702 bp in length
 29537 29636: gap of 100 bp
 29637 30324: contig of 688 bp in length
 30325 30424: gap of 100 bp
 30425 31130: contig of 706 bp in length
 31131 31230: gap of 100 bp
 31231 31910: contig of 680 bp in length
 31911 32010: gap of 100 bp
 32011 32691: contig of 681 bp in length
 32692 32791: gap of 100 bp
 32792 33482: contig of 691 bp in length
 33483 33582: gap of 100 bp
 33583 34274: contig of 692 bp in length
 34275 34374: gap of 100 bp
 34375 35081: contig of 707 bp in length
 35082 35181: gap of 100 bp
 35182 35861: contig of 680 bp in length
 35862 35961: gap of 100 bp
 35962 36660: contig of 699 bp in length
 36661 36760: gap of 100 bp
 36761 37447: contig of 687 bp in length
 37448 37547: gap of 100 bp
 37548 38243: contig of 696 bp in length
 38244 38343: gap of 100 bp
 38344 39034: contig of 691 bp in length
 39035 39134: gap of 100 bp
 39135 39813: contig of 679 bp in length
 39814 39913: gap of 100 bp
 39914 40597: contig of 684 bp in length
 40598 40697: gap of 100 bp
 40698 41392: contig of 695 bp in length
 41393 41492: gap of 100 bp
 41493 42190: contig of 698 bp in length
 42191 42290: gap of 100 bp
 42291 42967: contig of 677 bp in length
 42968 43067: gap of 100 bp
 43068 43736: contig of 669 bp in length
 43737 43836: gap of 100 bp
 43837 44525: contig of 689 bp in length
 44526 44625: gap of 100 bp
 44626 45306: contig of 681 bp in length
 45307 45406: gap of 100 bp
 45407 46111: contig of 705 bp in length
 46112 46211: gap of 100 bp
 46212 46848: contig of 637 bp in length
 46849 46948: gap of 100 bp
 46949 47639: contig of 691 bp in length
 47640 47739: gap of 100 bp

DEFINITION	REFERENCE
ACCESSION	AUTHORS
VERSION	TITLE
KEYWORDS	JOURNAL
SOURCE	REFERENCE
ORGANISM	AUTHORS
	TITLE
	JOURNAL
	FEATURES
	SOURCE

BASE COUNT ORIGIN	CDS
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Query Match	Best Local	Matches
39	G	QY
76	G	Db
99	A	QY
136	A	Db
159	C	QY
196	T	Db
219	T	QY
256	T	Db
279	A	QY
316	A	Db
339	A	QY
376	A	Db
399	T	QY
436	T	Db
459	C	QY
496	C	Db

TGGGCGCTTAGTGACCTTAGAAGTTTACCCATGCTCACTTCATCAATT 518
||||| ||||| ||||| | ||||| ||| | |||||
TGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCCTTATAAATC 555


```
Db 890 TGGCTTTCTGGAACAGTGTATCAACCCCTGCTCTCTATTTCTTTGGGAGATCACTTCA 949
Qy 936 GAGAGATGCTGATTAAGTTCAGACAATATCTCAAGTCCCTTACATCTTCCAGGACAT 995
Db 950 GGGACATGCTGATGAATCACTAGACACAACTTCAAACTCCCTTACATCTTTAGCAGAT 1009
Qy 996 GAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1048
Db 1010 GGGCTATGAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1063

RESULT 7
BC030948 1449 bp mRNA linear PRI 13-JUN-2002
LOCUS Homo sapiens, G protein-coupled receptor 91, clone MGC:32514
DEFINITION IMAGE:4594810, mRNA, complete cds.
ACCESSION BC030948
VERSION BC030948.1 GI:21410927
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1449)
Strausberg,R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 41 Row: e Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14780893.
Location/Qualifiers
1..1449
/organism="Homo sapiens"
/db_xref="LocusID:56670"
/db_xref="taxon:9606"
/clone="MGC:32514 IMAGE:4594810"
/tissue_type="kidney"
/clone_lib="NIH_MGC_75"
/lab_host="DH10B"
/note="vector: pDNR-LIB"
100..1104
/codon_start=1
/product="G protein-coupled receptor 91"
/protein_id="AAH30948.1"
/db_xref="GI:21410928"
/translation="MLGIMAWNATCKNWLAAEALEKYLISIFVGFVVGVLGNTIV
VYGFISLKNWNSNLYLENLSVDSLAFLCTPLMLIBSYANGNWIQDVLCISNRVYL
HANLYTSILEFTSIDRYLLIKYPFRELHLLQKEFALLSLAILWLVLTLELLPLPL
INPVITDNGTCNDFASSGDPNLYSMCLTLLGLFLPFLVCMCFYFKTALFLKORN
QVATLPLEKPLNLYIMAVIFSVLTPYHVMNRVIRASRLGWSKQYQCTQVINSF
YIVTRPLAFNSLVNIPVYFELLDGHRDMLMNLNRHNFKSLTSFSRWAHELLLSFREK
"

BASE COUNT 411 a 308 c 287 g 443 t
ORIGIN
```

```
Query Match 38.4%; Score 592.4; DB 9; Length 1449;
Best Local Similarity 75.3%; Pred. No. 6.5e-123;
Matches 764; Conservative 0; Mismatches 246; Indels 4; Gaps 2;

Qy 39 GCAGAAATGGCAGAGAAATTTATCTTGAGAAATTTGCTTGGCAACAGAGGCTATCTTGAATA 98
Db 107 GGATCATGGCAGTGAATGCAACTTGCAGAAATGCTTGGCAGCAGAGGCTGCCCTGGAA 166
Qy 99 AGTACTACTCTCTGCAATTTTATGCAATCGAGTTCATTTTGGAGTCTGCTGGGAATGCA 158
Db 167 AGTACTACTCTTCCATTTTATGGGATGAGTTCCTTGTGGGAGTCCCTTGGAAATACCA 226
Qy 159 CTGTGGTGTGGGTACCTCTCTCATGAAGAAGTCTGCAACAGCAGCAATCTCTATCTTT 218
Db 227 TGTGTTGTACGGCTACATCTCTCTGAAGAAGTCTGCAACAGCAGCAATATTTATCTCT 286
Qy 219 TTAACCTTTCCATCTCTGACTTTGCTTCTGTGACACCTTCCCATCTGATAAGAGTT 278
Db 287 TTAACCTCTCTCTGACTTTAGCTTTCTGTGACACCTCCCATCTGATAAGAGTT 346
Qy 279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATGAAGCAACCGGATATGCTTC 338
Db 347 ATGCCAATGGAAGAACTGGATATATGAGACGCTCTGTCATGAAGCAACCGGATATGCTTC 406
Qy 339 ACACCAACTCTACACCAAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398
Db 407 ATGCCAACTCTATACCAAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 466
Qy 399 TCATGAAGTACCCCTTCCGAGAAGCACTTTCTACAAAGAGAAATTTGCCATTTTAATCT 458
Db 467 TAAATAGTATCTTTCGGAACACCTCTCTGCAAAAGAAAGAGATTTGCTATTTTAATCT 526
Qy 459 CGCTGGCTGTCTGGCCCTTAGTGACCTTAGAAGTTCTACCATCTCTCTCTCTCTCTCTCTCT 518
Db 527 CTTTGGCCATTTGGGTTTGTAGTAACCTTAGAGTTACTACCATCTCTCTCTCTCTCTCTCTCT 586
Qy 519 CTGTCCCAAAAGAGAGAGAGAGTAACTGCACTGCACTTGTCAAGTCTTGGAAACCCCTCAAC 578
Db 587 CTGTATAACTGACAATGGCACCACCTGTATGATTTTGCAGTTCTGGAGACCCCACT 646
Qy 579 ACAATCTATTTACAGGCTCTGCTGCTGCTTGTGGGCTTCTCTCTCTCTCTCTCTCTCTCT 638
Db 647 ACAACCTCATTTACAGCATGTGTCTAAACCTGTGGGTTTCTTATCTCTCTCTCTCTCTCT 706
Qy 639 TGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 698
Db 707 TGTGTTTCTTTTATACAAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 766
Qy 699 CTGCCCTGCCACTGGACAACCCCAACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCT 758
Db 767 CTGCTCTGCCCTTTGAAAGCCTCTCAACTTGGTTCATCATGGCAGTGTGTAATCTCTCTCTG 826
Qy 759 TACTCTTACACCCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCTCGATGATTT 818
Db 827 TGTCTTTTACACCCCTATCATGCGCAATGTGAGGATCGCTTTCACGCTCTGCGGAGTT 886
Qy 819 G---GCCACAAGGATGTACACAGAAGGCCATCAAACTCTATATACACTCACACGCGCTC 875
Db 887 GGAAGCAGTATCAGTGCACACAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 946
Qy 876 TGGCCTTTCTGAACAGTGGCCATCAATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 935
Db 947 TGGCCTTTCTGAACAGTGTCTATCAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1006
Qy 936 GAGAGATGCTGATTAAGTTCAGACAATCTCAAGTCCCTTACATCTCTCTCTCTCTCTCTCTCT 995
Db 1007 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAACTCCCTTACATCTCTCTCTCTCTCT 1066
Qy 996 GAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1048
Db 1067 GGGCTCATGAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1120
```


Scott, G., Shen, H., Shooshtari, N., Sisson, I., Sodergren, E.,
 Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A.,
 Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C.,
 Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L.,
 Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S.,
 Warren, R., Washington, C., Watlington, S., Williams, G.,
 Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,
 Wu, Y. F., Zhou, J., Zorrilla, S., Naylor, S. L., Weinstein, G. and
 Gibbs, R.

Direct Submission
 Unpublished
 2 (bases 1 to 90343)
 Worley, K.C.

Direct Submission
 Submitted (23-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 90343)
 Worley, K.C.

Direct Submission
 Submitted (09-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
 gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
 entire insert of this clone. Overlapping regions of clones are only
 sequenced and submitted once, so the sequence for the remainder of
 the insert may be found in the record for the adjacent clones.
 Overlapping clones are noted at the beginning and end of the
 Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
 of a local database that includes entries from dbSTS, GDB, and
 local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
 unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
 EST and cDNA sequences. Genes demonstrate at least two exons
 flanked by consensus splice sites that maintained sequence
 continuity across the splice junctions. Sequences that are not
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
 standard of double strand coverage with a minimum of 2 clones and 2
 reads with no ambiguities or 2 chemistries with a minimum of 2
 clones and 3 reads with no ambiguities. If the sequence quality for
 a region does not meet this standard, it will be indicated in the
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
 standards - estimated error rate less than 1 per 10,000 bases.
 Reports of lowest quality individual bases and measures of base
 quality are listed below. Description of the metrics can be found
 at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

Location/Qualifiers
 1. .90343
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-3F11"
 991. .1106
 /rpt_family="MER45B"
 complement(1314. .1627)
 /rpt_family="AluXs"
 complement(2137. .2430)
 /rpt_family="AluY"

repeat_region
 complement(2568. .2741)
 /rpt_family="L1M4"
 repeat_region
 complement(2742. .3047)
 /rpt_family="AluXs"
 complement(3048. .3165)
 /rpt_family="L1M4"
 4735. .4865
 /rpt_family="FLAM_C"
 5657. .5762
 /rpt_family="L1MC/D"
 5906. .6237
 /rpt_family="L1R21B"
 6289. .6773
 /rpt_family="HERVPH21"
 complement(8725. .9597)
 /rpt_family="MER11D"
 12399. .12689
 /standard_name="136046"
 13774. .13816
 /rpt_family="Alu"
 13817. .13874
 /rpt_family="(TA)n"
 complement(15157. .15633)
 /rpt_family="L2"
 15706. .15747
 /rpt_family="AT_rich"
 16025. .16235
 /rpt_family="MIR"
 16560. .16682
 /rpt_family="L2"
 complement(16710. .17265)
 /rpt_family="L1R49"
 18077. .18368
 /rpt_family="AluXs"
 complement(18376. .18471)
 /rpt_family="L2"
 complement(18486. .18859)
 /rpt_family="MER57B"
 complement(20618. .20922)
 /rpt_family="AluXs"
 21337. .21363
 /rpt_family="AT_rich"
 22155. .22561
 /rpt_family="L1M4"
 complement(22608. .22659)
 /rpt_family="L1M4"
 22685. .23013
 /rpt_family="L1MB8"
 23103. .23399
 /rpt_family="AluSg"
 23500. .23973
 /rpt_family="L1MB3A"
 complement(24027. .24305)
 /rpt_family="L1MB1"
 24304. .24655
 /rpt_family="L1MB3A"
 24656. .24697
 /rpt_family="MADE1"
 25203. .25518
 /rpt_family="AluJo"
 25783. .25817
 /rpt_family="(TAA)n"
 26187. .26211
 /rpt_family="AT_rich"
 27014. .27030
 /rpt_family="AT_rich"
 complement(27031. .27316)
 /rpt_family="AluXs"
 27317. .27328
 /rpt_family="AT_rich"
 27574. .27615
 /rpt_family="(TAGA)n"
 28062. .28166
 STS


```
misc_feature 1339..2500
/note="assembly_fragment"
2601..3814
/note="assembly_fragment"
3915..5688
/note="assembly_fragment"
5789..7453
/note="assembly_fragment"
7554..8459
/note="assembly_fragment
clone_end:Sp6
vector_side:right"
8560..11388
/note="assembly_fragment"
11489..14580
/note="assembly_fragment"
14681..17562
/note="assembly_fragment"
17663..20871
/note="assembly_fragment"
20972..24226
/note="assembly_fragment"
24327..28458
/note="assembly_fragment"
28559..32034
/note="assembly_fragment"
32135..34919
/note="assembly_fragment"
35020..37828
/note="assembly_fragment"
37929..42243
/note="assembly_fragment"
42344..44474
/note="assembly_fragment
clone_end:r7"
vector_side:right"
44575..48293
/note="assembly_fragment"
48394..53132
/note="assembly_fragment"
53233..57872
/note="assembly_fragment"
57973..63120
/note="assembly_fragment"
63221..68799
/note="assembly_fragment"
68900..74014
/note="assembly_fragment"
74115..79049
/note="assembly_fragment"
79150..85094
/note="assembly_fragment"
85195..90567
/note="assembly_fragment"
90668..96258
/note="assembly_fragment"
96359..103095
/note="assembly_fragment"
103196..110330
/note="assembly_fragment"
110431..117484
/note="assembly_fragment"
117585..126459
/note="assembly_fragment"
126560..134924
/note="assembly_fragment"
135025..143468
/note="assembly_fragment"
143569..158144
/note="assembly_fragment"
46890 a 29145 c 29539 g 49189 t 3381 others
ORIGIN
```

```
Query Match 38.1%; Score 588.6; DB 2; Length 158144;
Best Local Similarity 75.4%; Pred. No. 5.4e-122;
Matches 759; Conservative 0; Mismatches 224; Indels 4; Gaps 2;

QY 46 GGCACAGAAATTTATCTGTGAGAAATGGTTGGCAACAGAGGCTATCTTGAATAAGTACTA 105
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142395 GGCATGAATGCAACTTGGCAACACTGGCTGGCAGCAGAGGCTGCCCTGGAATAAGTACTA 142336
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 106 CCTCTCTGCATTTTATGCAATCGAGTTCAATTTTGGAGTCTGCTTGGGAATCTACTGTGT 165
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142335 CCTTTCATTTTATGGATTGAGTTCTGTTGGGAGTCTTGGAAATACCATGTTGT 142276
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166 GTTCGGCTACTCTTCTGCAATGAAGAACTGGAACAGCAGCAATGCTATCTTTTAACTT 225
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142275 TTACGGCTACATCTCTCTGAAAGAACTGGAACAGCAGTAATATTTATCTTTAACTT 142216
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 226 TTCCATCTCTGACATTTGCTTTCCTGTCACCTTCCCATCTCTGATAAAGAGTTATGCCAA 285
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142215 CTCTGCTCTCTGACATTTAGCTTTTCTGTCACCTCCCTCCCATCTGTAAGGATTTATGCCAA 142156
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 286 TGATAAGGGGACCTTATGGAGATTTCTCTGTATAGCAACCGATATGCTTTCACACCAA 345
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142155 TGGAACTGGATATATGAGACGTGCTCTGCATTAAGCAACCGATATGCTTTCATGCCAA 142096
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 346 CCTCTACACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 405
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142095 CCTCTATACACGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 142036
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 406 GTACCTTTCCGAGAACACTTTCTACAAAGAGGAATTTGCCATTTTAACTCTGCTGGC 465
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142035 GTATCTTTCCGAGAACACTTTCTGCAAAAGAGAGTTTGCTATTTTAACTCTCTGGC 141976
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 466 TGCTGGGCTTTAGTGACCTTAGAAGTTCTTACCATGCTCTCTCTCTCTCTCTCTCTCT 525
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141975 CATTTGGGTTTTAGTAACTTTAGAGTTTACTACCATCTTCCCTCTTAACTCTCTGTAT 141916
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 526 AAAAGAGAGGCGAGTAACATGCACTGCTCAAGTTCTGGAAGTTCTGGAACCTGGAACAATCT 585
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141915 AACTGCAATGGCACCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 141856
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 586 CATTTACAGCCTCTGCTGACTTTTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 645
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141855 CATTTACAGCATGCTCTAACACTGTTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCT 141796
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 646 CTTCTACTACAGATGTTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCACTGCCCT 705
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141795 CTTTATTTACAAGATTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 141736
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 706 GCACCTGGACAAACCCCAACGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 765
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141735 GCCCTTGAAGAGCTCTCAACTGGTCTCATCTGGCAGTGTATCTCTCTCTCTCTCTCT 141676
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 766 CACACCTTATCATATCATGCGCAATTTAGGATCGCTCTCAGCCTCTGATAGTTG---GCC 822
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141675 TACACCTTATCAGCTCATCGGGAATGTGAGGATCGCTTTCAGGCTGGGAGTTGGAAGA 141616
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 823 ACAAGATGTACACAGAGGCGCAATCTATATACACTGACACGCGCTCTGGCTT 882
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141615 GTATCAGTGCATCGCTCGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 141556
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 883 TCTGAACAGTGCATCAATCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 942
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141555 TCTGAACAGTGCATCACTCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 141496
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 943 GCTGATTAGTAAAGTTACAGCAATCTTCAAGTCCCTTACATCTCTTACAGCATGAGTGC 1002
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141495 GCTGATGAATCAACTGAGACAACTTCAAAATCCCTTACATCTCTTACAGATGGGTCA 141436
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1003 TGGATGTCAGGCTTCTCACTCAGCCAAA-TGAGACACTTTGATAAACAG 1048
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141435 TGAACCTCTACTTTTCTTATTCAGAGAAAAGTGGGGCTTGTGAAACAG 141389
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 13	AF031897	1977 bp	mRNA	linear	VRT 06-JAN-1998
LOCUS	Meleagris gallopavo G protein coupled p2Y nucleotide receptor				
DEFINITION	(p2Y) mRNA, complete cds.				
ACCESSION	AF031897				
VERSION	AF031897.1	GI:2707255			
KEYWORDS	Meleagris gallopavo.				
SOURCE	Meleagris gallopavo.				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Meleagridae;				
REFERENCE	1 (bases 1 to 1977)				
AUTHORS	Boyer, J.L., Waldo, G.L. and Harden, T.K.				
TITLE	Molecular cloning and expression of an avian G protein-coupled p2Y receptor				
JOURNAL	Mol. Pharmacol. 52 (6), 928-934 (1997)				
MEDLINE	98086419				
PUBMED	9415702				
REFERENCE	2 (bases 1 to 1977)				
AUTHORS	Boyer, J.L., Waldo, G.L. and Harden, T.K.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-Oct-1997) Dept. of Pharmacology, University of North Carolina School of Medicine, CB# 7365 Mary Ellen Jones Building, Chapel Hill, NC 27599-7365, USA				
FEATURES	Location/Qualifiers				
source	1. 1977				
gene	/organism="Meleagris gallopavo"				
	/db_xref="taxon:9103"				
	/tissue_type="blood"				
	1. 1977				
CDS	/gene="tp2y"				
	374..1498				
	/gene="tp2y"				
	/codon_start=1				
	/product="G protein coupled p2Y nucleotide receptor"				
	/protein_id="AAC60339.1"				
	/db_xref="GI:2707256"				
	/translation="MDAPVRFSLAPWTPTPTPLGNGNTTAAAEAKVNEEFKILL PISYGVFVVGPLNSWAMWIFSVSRPNHNTYFNLAISDTLYVSLPLVYYA DRNNPFGVFKIVRFLEYANLYSILFLTCISVHRHIGCHIRLSUKWVKHARL ICGVWVTCILPLNLEFVTTSSKNDLCHDHTKPEEFHYHYSIMALLFQIP FLVIVVYCMARKLCRFPSPVPSVKRSIKMIIIVLTVFAICFVFPFHRTL YTSRYFADQCLTNIINFTYKTRPLASINSLDPLIFYMAGDKYRGLRRGAQRP RPVPTSLALVSPVSDSVVGVGSCNSRGMGTVWSRGQ"				
BASE COUNT	416 a 607 c 531 g 423 t				
ORIGIN					
Query Match	9.6%; Score 147.6; DB 5; Length 1977;				
Best Local Similarity	51.9%; Pred. No. 9.7e-23;				
Matches	436; Conservative 0; Mismatches 389; Indels 15; Gaps 4;				
QY	119	TATGCAATCGAGTATTTTGGACGCTGTTGGGAATGTCACGTGGTGTGGCTACCTC	178		
Db	515	TAGGCATCTCTCTGGTGGTCTGCGCCCTCACTTCCTGGGCCATGGATTTTGTG	574		
QY	179	TTCTGCATGAAGAACTGGACAGCAGCAATGTCTATCTTTTAAACCTTTCCATCTCTGAC	238		
Db	575	TCCAGATGAGGCTTGGGAATGCCACCACTACTTACATGTTCAACCTGGCTATCTGAC	634		
QY	239	TTTGTCTTCTGTCACCTTCCATCTCTGATTAAGAGTTATGCCAATG---ATAAGGG	295		
Db	635	ACACTCTATGTCCTCTCCCTCCCAACCTGGTCTACTATTATGCTGACCGCAACACTGG	694		
QY	296	ACATGAGAGATCTCTCTGTATAGCAACCGATATGCTCTCACACCAACCTCTACAC	355		
Db	695	CCCTCGGAAAGTGTCTGCAAGATCGTGGCGCTTCCCTCTTACGCCAACCTCTACAC	754		
QY	356	AGCATCTCTCTCTCACTTTCAATAGCATGGACCGCATCTCTGCTCATGAAGTACCTTTC	415		
Db	755	AGCATCTCTCTCTGACGTGCATCAGCTCCACCGCTACATGGGCATCTGGCCACCCATC	814		
QY	416	CGAAGAACACTTTTACAAAGAGAGGAATTTGCCATTTTAACTCTGCTGCTGTCTGGGCC	475		
Db	815	CGTCTCCTCAAGTGGGTGAAGCAAAACATGCACGACTCATCTGGGTGGCGCTGTGGCTG	874		
QY	476	TTAGTGAACCTTAGAAGTTTACCCATGCTTACCTTTTTCATCAATTTCTGTCCTCAAAAGAG	535		
Db	875	GTGTGTCACCATCTGCTCATCCCAACCTCATCTTCTGTCACCAACGCTCCCAAGACAC	934		
QY	536	GGCAGTAACATGCAATGCACTATGCAAGTTCTGGAACCCCTGAACAAATCTATTACAGC	595		
Db	935	AGCACCTTTTGGCATGACACCAACCAACCTGAGGAGTTCCGACCACTACGTCACCTACAG	994		
QY	596	CTCTGCTGACITTTGTTGGGCTTCTTAATCTCTCTCTCTGTCATGTCCTTCTTCTACTAC	655		
Db	995	TCTCTCATGTCGCTCTCTTCTGGTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1054		
QY	656	AAGATGGTAGTCTTCTTAAAGAGGAGGAGC-----CAGCAGCAAGCAACTGCCCTGC	707		
Db	1055	CTGATGGCCAAGAGGCTCTGCAAGCGCAGCTTCCCGCAGCCCGCCCGCTGCCCTCC	1114		
QY	708	CAGTGACAAACCCCAAC--GCCTGGTGGHCTGGGGTGTGTGATCTTCTCTATATCTTTC	766		
Db	1115	TACAAGAGCGTCCCATCAAGATGATCATCTGCTCACCGCTCTCGCCATTTGCTTT	1174		
QY	767	ACACCTTATCATATCATGCGCAATTTGAGGATCGCTCACGCTCGATAGTTGGCCACAA	826		
Db	1175	GTGCTCTCCACATCACCGCGACCTCTACTACACTCTCGC---TACTTCCAAGCTGAT	1231		
QY	827	GGATGTACAGAGAGGCCATCAATCTATATACACTGACACGCGCTCTGGCCTTTCTG	886		
Db	1232	TGTACAGCCTCAACATCATCAACTTCACTTACAAGATCACGCGACCTCGCCAGCATC	1291		
QY	887	AACAGTGCCATCAATCCCATCTTCTACTTCTCTGAGGAGACCATTACAGAGATGCTG	946		
Db	1292	AACAGTGCTGGACCCCATCTGTACTTCTGCTGGGAGCAAGTACCGGGGCGGCTG	1351		
RESULT 14	AC101335	69462 bp	DNA	linear	HTG 23-NOV-2001
LOCUS	Mus musculus clone RP23-109K24, LOW-PASS SEQUENCE SAMPLING.				
DEFINITION	AC101335				
ACCESSION	AC101335.1	GI:17060110			
VERSION	HTG; HTGS_PHASE0.				
KEYWORDS	Mus musculus.				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 69462)				
AUTHORS	Birren, B., Linton, L., Nusbaum, C. and Lander, E.				
TITLE	Unpublished				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 69462)				
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckhagalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J., Chararo, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Glende, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lebecky, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McSheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, M., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupp, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,				

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L16470
Center clone name: 109_K_24

* NOTE: This record contains 85 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
705 804 : contig of 704 bp in length
805 1533 : contig of 729 bp in length
1534 1633 : gap of 100 bp
1634 2366 : contig of 733 bp in length
2367 2466 : gap of 100 bp
2467 3194 : contig of 728 bp in length
3195 3294 : gap of 100 bp
3295 3987 : contig of 693 bp in length
3988 4087 : gap of 100 bp
4088 4826 : contig of 739 bp in length
4827 4926 : gap of 100 bp
4927 5647 : contig of 721 bp in length
5648 5747 : gap of 100 bp
5748 6477 : contig of 730 bp in length
6478 6577 : gap of 100 bp
6578 7292 : contig of 715 bp in length
7293 7392 : gap of 100 bp
7393 8113 : contig of 721 bp in length
8114 8213 : gap of 100 bp
8214 8933 : contig of 720 bp in length
8934 9033 : gap of 100 bp
9034 9758 : contig of 725 bp in length
9759 9858 : gap of 100 bp
9859 10651 : contig of 693 bp in length
10652 10651 : gap of 100 bp
10652 11386 : contig of 735 bp in length
11387 11486 : gap of 100 bp
11487 12223 : contig of 737 bp in length
12224 12323 : gap of 100 bp
12324 13057 : contig of 734 bp in length
13058 13157 : gap of 100 bp
13158 13845 : contig of 688 bp in length
13846 13945 : gap of 100 bp
13946 14669 : contig of 724 bp in length
14670 14769 : gap of 100 bp
14770 15482 : contig of 713 bp in length
15483 15582 : gap of 100 bp
15583 16293 : contig of 711 bp in length
16294 16393 : gap of 100 bp
16394 17089 : contig of 696 bp in length
17090 17189 : gap of 100 bp
17190 17914 : contig of 725 bp in length
17915 18014 : gap of 100 bp
18015 18576 : contig of 562 bp in length

18577 18676 : gap of 100 bp
18677 19389 : contig of 713 bp in length
19390 19489 : gap of 100 bp
19490 20233 : contig of 744 bp in length
20234 20333 : gap of 100 bp
20334 21062 : contig of 729 bp in length
21063 21162 : gap of 100 bp
21163 21865 : contig of 703 bp in length
21866 21965 : gap of 100 bp
21966 22694 : contig of 729 bp in length
22695 22794 : gap of 100 bp
22795 23520 : contig of 726 bp in length
23521 23620 : gap of 100 bp
23621 24337 : contig of 717 bp in length
24338 24437 : gap of 100 bp
24438 25149 : contig of 712 bp in length
25150 25249 : gap of 100 bp
25250 25905 : contig of 656 bp in length
25906 26005 : gap of 100 bp
26006 26738 : contig of 733 bp in length
26739 26838 : gap of 100 bp
26839 27572 : contig of 734 bp in length
27573 27672 : gap of 100 bp
27673 28400 : contig of 728 bp in length
28401 28500 : gap of 100 bp
28501 29196 : contig of 696 bp in length
29197 29296 : gap of 100 bp
29297 30013 : contig of 717 bp in length
30014 30113 : gap of 100 bp
30114 30821 : contig of 708 bp in length
30822 30921 : gap of 100 bp
30922 31642 : contig of 721 bp in length
31643 31742 : gap of 100 bp
31743 32459 : contig of 717 bp in length
32460 32559 : gap of 100 bp
32560 33291 : contig of 732 bp in length
33292 33391 : gap of 100 bp
33392 34135 : contig of 744 bp in length
34136 34235 : gap of 100 bp
34236 34946 : contig of 711 bp in length
34947 35046 : gap of 100 bp
35047 35775 : contig of 729 bp in length
35776 35875 : gap of 100 bp
35876 36605 : contig of 730 bp in length
36606 36705 : gap of 100 bp
36706 37439 : contig of 734 bp in length
37440 37539 : gap of 100 bp
37540 38270 : contig of 731 bp in length
38271 38370 : gap of 100 bp
38371 39096 : contig of 726 bp in length
39097 39196 : gap of 100 bp
39197 39917 : contig of 721 bp in length
39918 40017 : gap of 100 bp
40018 40726 : contig of 709 bp in length
40727 40826 : gap of 100 bp
40827 41544 : contig of 718 bp in length
41545 41644 : gap of 100 bp
41646 42332 : contig of 688 bp in length
42333 42432 : gap of 100 bp
42433 43168 : contig of 736 bp in length
43169 43268 : gap of 100 bp
43269 44006 : contig of 738 bp in length
44007 44106 : gap of 100 bp
44107 44838 : contig of 732 bp in length
44839 44938 : gap of 100 bp
44939 45673 : contig of 735 bp in length
45674 45773 : gap of 100 bp
45774 46500 : contig of 727 bp in length
46501 46600 : gap of 100 bp
46601 47316 : contig of 716 bp in length
47317 47416 : gap of 100 bp
47417 48119 : contig of 703 bp in length
48120 48219 : gap of 100 bp

* 48220 48933: contig of 714 bp in length
* 48934 49033: gap of 100 bp
* 49034 49728: contig of 695 bp in length
* 49729 49828: gap of 100 bp
* 49829 50547: contig of 719 bp in length
* 50548 50647: gap of 100 bp
* 50648 51365: contig of 718 bp in length
* 51366 51465: gap of 100 bp
* 51466 52182: contig of 717 bp in length
* 52183 52282: gap of 100 bp
* 52283 53028: contig of 746 bp in length
* 53029 53128: gap of 100 bp
* 53129 53847: contig of 719 bp in length
* 53848 53947: gap of 100 bp
* 53948 54680: contig of 733 bp in length
* 54681 54780: gap of 100 bp
* 54781 55433: contig of 703 bp in length
* 55484 55583: gap of 100 bp
* 55584 56316: contig of 733 bp in length
* 56317 56416: gap of 100 bp

Query Match 9.4% Score 145.2; DB 2; Length 69462;
Best Local Similarity 98.0%; Pred. No. 3.8e-22;
Matches 147; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1394 AAGCGTGTCTGAAATAGTCTTATTTTATCTGTGAATATTAATAATTTATGTGAAAAA 1453
Db 4115 AATCTGTCTGAAATAGTCTTATTTTATCTGTGAATATTAATAATTTATGTGAAAAA 4174
Qy 1454 TGAATATAATTCATGACACATAGATTTCTTATTTGAAATATATATTTCTTGTAAAAA 1513
Db 4175 TGAATATAATTCATGACACATAGATTTCTTATTTGAAATATATATTTCTTGTAAAAA 4234
Qy 1514 ATAACTGCTGCGCTAAATTAATCAATATA 1543
Db 4235 ATAACTGCTGCGCTAAATTAATCAATATA 4264

RESULT 15
AX148186 1014 bp DNA linear PAT 08-JUN-2001
LOCUS AX148186
DEFINITION Sequence 27 from Patent WO0136471.
ACCESSION AX148186
VERSION AX148186.1 GI:14347086
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS Chen,R., Dang,H.T. and Lowitz,K.P.
TITLE Endogenous and non-endogenous versions of human g protein-coupled receptors
JOURNAL Patent: WO 0136471-A 27 25-MAY-2001;
ARENA Pharmaceuticals, Inc. (US)
FEATURES
source 1..1014
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 258 a 263 c 189 g 304 t
ORIGIN

Query Match 8.2% Score 126.6; DB 6; Length 1014;
Best Local Similarity 49.9%; Pred. No. 5.1e-18;
Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;
Qy 60 CTTGTGAGAAATGGTTGGCAACAGAGGCTATCTTTGAATAAGTACTACTCTCTGCAATTT 119
Db 59 CTTTGGAAATGCAGTGAATAACATCCCACTCAAGATGCATACCTCCCTGTTATTT 118
Qy 120 ATGCAATCGAGTTCAATTTTGGAGTCTTGGGAATGCACGTGGTCTCGGCTACCTCT 179
Db 119 ATGGCAATATCTCTCGTGGGATTTCCAGCAATGCAGTAGTATATATCCACTTACATTT 178

Qy 180 TCTGCATGAAGAACTGGAACAGCAATGTCTATCTTTTAACTTTTCCATCTCTGACT 239
Db 179 TCAAAATGAGACCTTGGAGAGCAGCACCATCATTAATGCTGAACCTGGCTGCACAGATC 238
Qy 240 TTGCTTTCCCTGTGACACCTTCCCATCTGATAAAGATTTATGCCAAT---GATAAGGGA 296
Db 239 TGCTGTATCTGACCAAGCTCCCTTCTCTGATTCACCTACTATGCCAGTGGCAAACTGGA 298
Qy 297 CCTATGGAGATGTTCTCTGTATATAAGCAACCGATATGTGCTTACACCACCTCTACACCA 356
Db 299 TCTTTGGAGATTTCATGTGTAAAGTTTATCCGCTTCAGCTTCCATTTCAACCTGTATAGCA 358
Qy 357 GATCCTCTTCCCTCACTTTTCATAGCATGACGACGATATCTGCTCATGAAGTACCCTTTCC 416
Db 359 GCATCCTCTTCCCTCACTTTTCAGCATCTTCGGCTACTGTGTGATCATTCACCCAATGA 418
Qy 417 GAGAACACTTTCTACAAAAGAGAAATTTGGCAATTTAATCTCGCTGGCTGTCTGGGCT 476
Db 419 GCTGCTTTTCCCAITCACAAAACCTCGATGTGCAGTTGTAGCTGTGTGTGTGATCA 478
Qy 477 TAGTGACCTTAGAAGTCTTACCCATGCTCCTCATCAATTTCTGCCAAAAGAGAGG 536
Db 479 TTTCACCTGGTAGTGTCAATTCGATGACCTTCTTGATCATCAACCAACAGAGCAACA 538
Qy 537 GCAGTAACCTGCATCGACTATGCAAGTTCTGGAAACCTGAAACACAATCTCATTTACAGCC 596
Db 539 GATCAGCTGTCTCGACCTCACCAGTTCCG-----ATCAACTCAATATAATTAAGTGT 592
Qy 597 TCTGCTGACTTTGTTGGGCTTCCCTAAATTCCTCTCTCTGTGTGTGTGTCTTCTACTACA 656
Db 593 ACAACCTGATTTTGACTGCAACTACTTTCTGCTCTCCCTTGGTGTGTGTGTGTGTGTGT 652
Qy 657 AGATGCTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACA 716
Db 653 ATACCAGATTTATCCACACTCTGACCCATGGACTGCAAACTGACAGCTGCCCTTAAGCAGA 712
Qy 717 AACCCCAACGCTGTGTGTGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 776
Db 713 AAGCAGGAAGGCTTAACCAATTTCTGCTACTCTTGCATTTTACGTATGTTTTTACCCCTGC 772
Qy 777 ATATCATGCGCAATTTGAGGATCGCCTCAGGCTG 811
Db 773 ATATCTTGAGGGTCAATTCGGATCGAATCTCGGCTG 807

Search completed: April 2, 2003, 14:32:25
Job time : 5026 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 2, 2003, 10:58:41 ; Search time 393 seconds
(without alignments)
8841.815 Million cell updates/sec

Title: US-09-891-138a-1

Perfect-score: 1543

Sequence: 1 gcccctggcagagtttctg.....tgcctaaataaatcaatata 1543

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1-Geneseq.101002.*

1: -/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1543	100.0	1543	24	ABK12957		DNA sequence of mouse G-protein coupled receptor TGR18
2	592.4	38.4	1005	21	AA001135		Human orphan G protein
3	592.4	38.4	1005	21	AA046036		Human G-protein co
4	592.4	38.4	1436	24	ABU90790		Human polynucleoti
5	592.4	38.4	1542	24	AA024958		Human G-protein co
6	589.2	38.2	1996	18	AA071900		Human purinergic r
7	587.2	38.1	1428	18	AA075146		Human ATP receptor
8	580.4	37.6	1385	21	AA081122		Human secreted pro
9	126.6	8.2	1011	24	AA043942		Human G-protein-co

10	126.6	8.2	1014	22	AA079948	Human cDNA encodin
11	126.6	8.2	1014	24	ABN85630	Human P2Y-like rec
12	126.6	8.2	1014	24	ABK11381	Human DNA encoding
13	126.6	8.2	1014	24	AAK98323	Human purinergic-r
14	126.6	8.2	1288	24	ABL56197	Human P2Y1-li enco
15	126.6	8.2	1729	22	AA08362	Human cDNA encodin
16	126.6	8.2	1729	23	ABV24026	Human prostate exp
17	126.6	8.2	1729	23	ABV25767	Human prostate exp
18	126.6	8.2	1729	23	ABV29909	Human prostate exp
19	126.6	8.2	1729	23	ABV30024	Human prostate exp
20	126.6	8.2	5435	24	ABL56198	Human P2Y1-li enco
21	126.6	8.2	9905	24	AAK98324	Human purinergic-r
22	125	8.1	1014	24	ABQ78847	Human G-protein co
23	125	8.1	1014	24	AA034278	Human G-protein co
24	125	8.1	1014	24	AA026370	Human AXOR89 (G-pr
25	119.2	7.7	1020	22	AAH51011	Human G-protein co
26	119	7.7	1313	22	AAK52430	Human polynucleoti
27	112.6	7.3	740	23	ABV15662	Human prostate exp
28	104	6.7	1076	24	AA029667	Human G-protein co
29	103.2	6.7	6721	24	AA018600	Purinergic recepto
30	100.8	6.5	6721	24	AA018599	Human G-protein-co
31	97.8	6.3	1002	22	AAH73516	Human purinergic r
32	97.8	6.3	3143	24	AA017746	Human genomic clon
33	96	6.2	831	24	ABN85629	Human P2Y-like rec
34	90	5.8	539	23	ABV39127	Human prostate exp
35	90	5.8	539	23	ABV45465	Human prostate exp
36	90	5.8	1063	24	AA098135	Human DNA for pote
37	88.4	5.7	1429	18	AA074321	Human P2Y4 recepto
38	87	5.6	1163	22	AA040981	Human purinergic r
39	85.4	5.5	2025	22	AA040980	Human purinergic r
40	84.8	5.5	2796	24	ABQ91998	Human NF-kB activa
41	83.8	5.4	1113	23	AB019810	Non-endogenous hum
42	83	5.4	850	24	AA017747	Human P2Y-like G p
43	83	5.4	2138	24	AA032937	Mus musculus GPCR
44	83	5.4	2138	24	AA032934	Mus musculus GPCR
45	80	5.2	1068	23	AB197979	Non-endogenous hum

ALIGNMENTS

RESULT 1
ABK12957
ID ABK12957 standard; DNA; 1543 bp.
XX
AC ABK12957;
XX

DT 09-APR-2002 (first entry)

XX DNA sequence of mouse G-protein coupled receptor TGR18 gene.

DE Mouse; G-protein coupled; receptor; GPCR; TGR18; kidney disease;
KW signal transduction modulator; cerebral cavernous malformation;
KW hyperlipidemia; obesity; dyslexia; cardiac myxoma; renal failure;
KW nephritis; hypertension; liver disease; cirrhosis; blood disorder;
KW spleen-associated disorder; immune disorder; gene; ds.

OS Mus sp.

XX Key Location/Qualifiers

FT CDS 44..997

FT /*tag= a

FT /product= "Mouse G-protein coupled receptor TGR18"

XX WO200200719-A2.

XX 03-JAN-2002.

XX 25-JUN-2001; 2001WO-US20363.

XX 23-JUN-2000; 2000US-213461P.

XX (TULA-) TULARIK INC.

RESULT 2

AD001135
 ID AAD01135 standard; cDNA; 1005 BP.
 AC AAD01135;
 XX
 XX 02-NOV-2000 (first entry)
 XX
 DE Human orphan G protein-coupled receptor hCHN10 cDNA.
 XX
 KW Human; orphan G protein-coupled receptor; GPCR; hCHN10; drug screening;
 transmembrane receptor; expressed sequence tag; EST; signal cascade; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1005
 FT /*tag= a
 FT /product= "hCHN10"
 FT /note= "Human orphan G protein-coupled receptor"

WO200031258-A2.
 XX
 XX
 XX
 PD 02-JUN-2000.
 XX
 PF 13-OCT-1999; 99MO-US23687.
 XX
 PR 20-NOV-1998; 98US-0109213.
 PR 16-FEB-1999; 99US-0120416.
 PR 26-FEB-1999; 99US-0121852.
 PR 12-MAR-1999; 99US-0123946.
 PR 12-MAR-1999; 99US-0123949.
 PR 28-MAY-1999; 99US-0136436.
 PR 28-MAY-1999; 99US-0136437.
 PR 28-MAY-1999; 99US-0136439.
 PR 28-MAY-1999; 99US-0136567.
 PR 28-MAY-1999; 99US-0137127.
 PR 28-MAY-1999; 99US-0137131.
 PR 29-JUN-1999; 99US-014448.
 PR 29-SEP-1999; 99US-0156555.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156634.
 PR 01-OCT-1999; 99US-0157280.
 PR 01-OCT-1999; 99US-0157281.
 PR 01-OCT-1999; 99US-0157282.
 PR 01-OCT-1999; 99US-0157293.
 PR 01-OCT-1999; 99US-0157294.
 PR 12-OCT-1999; 99US-0416760.
 PR 12-OCT-1999; 99US-0417044.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Chen R, Dang HT, Liaw CW, Lin I;
 XX
 DR WPI; 2000-400068/34.
 DR P-PSDB; AAY71308.
 XX
 PT Novel human orphan G protein-coupled receptors and the encoding cDNAs
 for use in the identification of G protein-coupled receptor agonists -
 XX
 PS Claim 69; Page 86; 102pp; English.
 XX
 CC The present sequence is a cDNA encoding hCHN10, an endogenous human
 orphan G protein-coupled receptor (GPCR), expressed in kidney and
 thyroid. The hCHN10 cDNA was identified using the human EST (expressed
 sequence tag) 1365839 as a probe. The orphan GPCR of the invention, like
 all GPCRs has seven transmembrane alpha helices with an extracellular
 N-terminus and an intracellular C-terminus. However, no endogenous
 ligands has yet been identified for the proteins of the invention. The
 orphan GPCRs may be used in the identification of their endogenous
 ligands, and to screen potential GPCR agonists and antagonists for use as
 pharmaceutical agents. The proteins may also be used in the study of

CC GPCR-mediated signalling cascades, and to elucidate their precise role in
 normal and diseased human conditions. Nucleic acid encoding human orphan
 GPCRs may be used for tissue localisation expression analysis to provide
 CC information about their function in healthy and pathological states.
 XX
 SQ Sequence 1005 BP; 248 A; 236 C; 196 G; 325 T; 0 other;

Query Match 38.4%; Score 592.4; DB 21; Length 1005;
 Best Local Similarity 75.5%; Pred. No. 1e-139;
 Matches 750; Conservative 0; Mismatches 241; Indels 3; Gaps 1;

QY 39 GCAGAATGGCACAGAATTTATCTTGAGAAATTTGGTTGGCAACAGAGGCTATCTTGAATA 98
 DB 8 GGATCATGGCATGGAATGCAACTTGCAAAACTGGCTGGCAGCAGGCTGCCCTGGAAA 67
 QY 99 AGTACTACCTCTCTGCAATTTTATGCAATCGAGTTTCATTTTGGACTGCTTGGGAATGTCA 158
 DB 68 AGTACTACCTTTCCATTTTATGGGATTTGAGTTCTGTTGGGAGTCTCTTGAATAACCA 127
 QY 159 CTGTGGTGTTCGGCTACCTCTTCTGCAAGAACTGGAACAGCAGCAATGCTATCTTT 218
 DB 128 TTGTTGTTTACGGCTACATCTTCTCTGAGAACTGGAACAGCAGTAATATTTATCTCT 187
 QY 219 TTAACCTTTCCATCTCTGACTTTGCTTCTGTCACCTTCCCATCTCTGATAAAGATT 278
 DB 188 TTAACCTCTCTCTGCTGACTTTTCTGTGACCTTCCCATCTCTGATAAAGATT 247
 QY 279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGATAAGCAACCGATATGCTTC 338
 DB 248 ATGCCAATGGAACCTGGATATATGGAGAGCTGCTCTGCAAGCAACCGATATGCTTC 307
 QY 339 ACACCAACCTCTACACAGCAGCCTCTTCTTCCACCTTTTCATTTAGCATGACCGATATCG 398
 DB 308 ATGCCAACCTCTATACCAGCATCTCTTCTCACCTTTTATCAGCATAGATGATCTTGA 367
 QY 399 TCATGAAGTACCTTTCCGGAACACTTTCTACAAAAGAGAAATTTGCCATTTTAACTCT 458
 DB 368 TAAATTAAGTATCCTTTCCGGAACACTTCTGCAAAAGAGAGTTTCTATTTTAACTCT 427
 QY 459 CGCTGGCTGCTCGGGCCCTTAGTGACCTTAGAAGTTCTACCCATGCTCATCTTTCATCA 518
 DB 428 CTTGGCCATTTGGGCTTTTAGTAAGCTTAGAGTTTACTACCCATCTTCCCTTTAAATC 487
 QY 519 CTGTCCCAAAAGAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAACCTCTGAAC 578
 DB 488 CTGTTTAACTGACAATGSCACCTCTGTAATGATTTTGCAGTTCTGGAGACCCCACT 547
 QY 579 ACAATCTCATTTACAGCCTCTGCTGAGCTTTGTTGGGCTTCTTAATTCCTCTCTGTGA 638
 DB 548 ACAACCTCATTTACAGCATGTCTTACACACTGTTGGGTTCTTATTTCTCTTTTGTGA 607
 QY 639 TGTGCTTCTCTACTACAGATGGTAGTCTTCTTAAAGAGAGGAGCCAGCAGCAAGCAA 698
 DB 608 TGTGTTCTTTTATACAGATGCTCTCTTCTTAAAGAGAGGAAATAGCAGGTTCTCTA 667
 QY 699 CTGCCCTGCCACTGGCAAAACCCCAAGCCCTGGTGGCTCTGCTGATCTTCTCTA 758
 DB 668 CTGCTCTGCCCTTGAAGAGCTCTCAACTTGGTCATCATGGCAGTGGTATCTTCTCTG 727
 QY 759 TACTCTTCACACCTTATCATATCATGCGCAATTTGAGGATCGGCTCACGCTCGATGATT 818
 DB 728 TGTCTTTTACACCTTATCAGTCTATGCGGAATGTGAGGATCGCTTACGCTCGGAGTT 787
 QY 819 G---GCCAAGAGGTGTACAGAAAGCCCATCAATCTATATACACTACACAGGCTCT 875
 DB 788 GGAAGCAGTATCAGTGCACTCAGTCTCATCAACTCTCTTTTACATGTGACACGGCTT 847
 QY 876 TGGCCTTTCTGAACAGTGGCCATCAATCCCATCTTCTACTCTTCATCGGAGACCATACA 935
 DB 848 TGGCCTTTCTGAACAGTGTCAACCCCTGCTCTCTATTTTCTTTTGGGAGATCATCTCA 907
 QY 936 GAGAGATGCTTGAATAGTAACTTACAGCAATPACTTCAAGTCCCTTACATCTCTCAGGACAT 995

QY 996 GAGTGTGGGATGACAGGCTTTCACTACGCCAAAA 1029
 DB III III I III III III III
 DB 968 GGGCTCATGAAGTCTCTACTTTTCATTCAGAGAAA 1001

RESULT 4

ABL90790
 ID ABL90790 standard; cDNA; 1436 BP.

XX ABL90790;

XX 24-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 1352.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; gene; ss.

XX Homo sapiens.

XX WO200190304-A2.

XX 29-NOV-2001.

XX 18-MAY-2001; 2001WO-US16450.

XX 19-MAY-2000; 2000US-205515P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-122018/16.

DR P-PSDB; ABB90381.

PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders.

PS Claim 4; SEQ ID NO 1352; 2081pp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB9040-ABB9044) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1436 BP; 397 A; 309 C; 289 G; 441 T; 0 other;

Query Match 38.4%; Score 592.4; DB 24; Length 1436;

Best Local Similarity 75.3%; Pred. No. 1.2e-139;
 Matches 764; Conservative 0; Mismatches 246; Indels 4; Gaps 2;

QY 39 GCAGAAATGCACAGAAATTTATCTGTGAGAAATGGTTGGCAACAGAGGCTATCTTGAATA 98
 I I IIIII IIII I III I III III III III III III III III III III III

DB 100 GGATCATGGCATGGAATGCAACTTGC AAAAAGCTGGTGGCAGCAGAGGCTGCCCTGGAAA 159
 QY 99 AGTACTACCTCTCTGCAATTTATGCAATCGAGTTCAATTTTGGACTGCTTTGGAAATGTCA 158
 DB IIIII III I IIIII III IIIII III III III III III III III III III III
 DB 160 AGTACTACCTTTCCCATTTTATGGAGTTGAGTTGCTGTGGAGTCCCTTGGAAATACCA 219
 QY 159 CTGTGGTGTTCGGCTACCTCTCTGCAATGGAAGAACTGGAACAGCAGCAATGCTATCTTT 218
 DB III III I IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 220 TTGTTGTTTACGGCTACATCTCTCTGCAAGAACTGGGAACAGCAATAATTTATCTCT 279
 QY 219 TTAACCTTTCCATCTCTGACTTTGCTTTCTGTGACCCCTTCCCATCTCTGATAAAGAGTT 278
 DB IIIII III IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 280 TTAACCTCTCTGCTCTGACTTTGCTTTCTGTGACCCCTTCCCATCTCTGATAAAGAGTT 339
 QY 279 ATGCCAATGATAGGGGACCTATGAGATGTTCTCTGTATATAAGCAACCAATATGCTTTC 338
 DB IIIII III III IIIII III IIIII III IIIII III IIIII III IIIII III IIIII
 DB 340 ATGCCAATGGAACCTGGATATATGAGAGCTGCTCTGCAATAAGCAACCAATATGCTTTC 399
 QY 339 ACACCAACCTCTACACGACGATCTCTCTCTGACCTTTTCAATAGCATGGACCAATATCTGC 398
 DB IIIII III IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 400 ATGCCAACCTCTATACCACGATCTCTTTCTCATTATATCAGCATAGATCGATCTTGA 459
 QY 399 TCATGAAGTACCCCTTCCGAGAACACTTTCTACAAAAGAGAAATTTGCCATTTTAATCT 458
 DB III III I IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 460 TAAITTAAGTATCCTTTCCGAGAACACCTCTGCAAAAAGAAAGAGTTTGTCTATTTTAATCT 519
 QY 459 CGCTGGCTGTCTGGGCTTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTTCATCAATT 518
 DB III III I IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 520 CCTTGGCCATTTGGGTTTGTAGTAACCTTAGAGTTTACTACCCATCTTCCCTTATAAATC 579
 QY 519 CTGTCCCAAAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGCAAAACCTCTGAAC 578
 DB III III I IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 580 CTGTTATAACTGACAAATGGCACCCCTGTATGATTTTTCGAGTTCTGGAGACCCCACT 639
 QY 579 ACAATCTCAATTTACAGCCTCTGCTGACTTTTGGGGTCTCTTAATCTCTCTCTGTGGA 638
 DB IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 640 ACAACCTCAATTTACAGCATGTCTTAACACTGTTGGGGTCTCTTATCTCTCTTTTGTGA 699
 QY 639 TGTGCTTCTCTACTACAGATGTTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
 DB III III I IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 700 TGTGTTTCTTTTATTACAAGATGCTCTCTTCCCTTAAGCAGAGGAAATAGGCAGGTGCTA 759
 QY 699 CTGCGCTGGCACTGGACAAACCCCAACGCTGGTGGTCTCTGGCGGTGTGTGATCTTCTCTA 758
 DB III III I IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 760 CTGCTCTGCCCTTGAAGGCTCTCAACTTGGTTCATCATGGCAGTGTGTAATCTTCTCTG 819
 QY 759 TACTCTTCAACCCCTATCATATCATGCGCAATTTGAGGATCGCTCAGCGCTGATAGTT 818
 DB III III I IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 820 TGCTTTTACACCCCTATCAGCTCATGCGGAATGTGAGGATCGCTTCAGCGCTGGGAGTT 879
 QY 819 G---GCCACAAGGATGTACACAGAGGCCATCAAAATCTATATACACACTGACACGGGCTC 875
 DB III III I IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 880 GGAAGCAGTATCAGTGGCACTCAGGTGCTCATCACTCTTTTACATCTGTGACACGGGCTT 939
 QY 876 TGGCCCTTTCTGAACAGTGGCAATCAATCCCATCTTCTACTTCCCTCATGGGAGACCATTACA 935
 DB III III I IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 940 TGGCCCTTCTGAACAGTGTCTCAACCCCTGTCTTCTATTTTCTTTTGGGAGATCATTCA 999
 QY 936 GAGAGATCTGATTAGTAAAGTTTACAGACAATCTTCAAGTCCCTTACATCTCTCAGGACAT 995
 DB III III I IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 1000 GGGACATGCTGATGAATCAACTGAGACACAACCTTCAATCCCTTACATCTTTAGAGAT 1059
 QY 996 GAGTGTGGATGAGGTTCTTCACTACGCCAAAA-TGAGACACTTTGATAAACAG 1048
 DB III III I IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 DB 1060 GGGCTCATGAAGTCTCTACTTTTCAATTCAGAGAAAAGTGGGGGCTTGTGAAACAG 1113

RESULT 5

AAD24958

ID AAD24958 standard; cDNA; 1542 BP.

XX AAD24958;

XX AAD24958;

DT 12-MAR-2002 (first entry)

XX Human G-protein coupled receptor-3 (GCREC-3) cDNA.

DE Human: G-protein coupled receptor-3; GCREC-3; therapy: cancer; stroke;

XX cell proliferative disorder; neurological; epilepsy; Parkinson's disease;

KW Alzheimer's disease; inflammation; thyroiditis; haemolytic anaemia; AIDS;

KW Acquired Immune Deficiency Syndrome; dementia; neutropenic; cholelithiasis;

KW Multiple sclerosis; atherosclerosis; angina pectoris; gastroenteritis;

KW diabetes; ulcer; viral infection; immunosuppressive; ss.

XX Homo sapiens.

XX Location/Qualifiers

PH 63..1202

FT /*tag= a

FT /product= "Human GCREC-3 protein"

FT WO200198351-A2.

XX 27-DEC-2001.

XX 15-JUN-2001; 2001WO-US19275.

XX 16-JUN-2000; 2000US-212483P.

XX 22-JUN-2000; 2000US-213954P.

XX 29-JUN-2000; 2000US-215209P.

XX 07-JUL-2000; 2000US-216595P.

XX 14-JUL-2000; 2000US-218936P.

XX 19-JUL-2000; 2000US-219154P.

XX 21-JUL-2000; 2000US-220141P.

XX (JNCY-) INCYTE GENOMICS INC.

XX Lal P, Baughn MR, Hafalia AJA, Nguyen DB, Gandhi AR, Kallick DA;

XX Griffin JA, Yue H, Khan FA, Patterson C, Lu DAM, Tribouley CM;

XX Lu Y, Walia NK, Grail R, Yang J, Ramkumar J, Au-Young J;

XX Elliott VS, Hernandez R, Walsh RT, Borowsky ML, Thornton M, He A;

XX WPI: 2002-075627/10.

XX P-PSDB: AAE15633.

XX Isolated human G-protein coupled receptor polypeptides and the use of

XX these sequences in the diagnosis, treatment and prevention of diseases

XX and in the assessment of exogenous compounds on the expression of the

XX receptors -

XX Claim 11; Page 133; 143pp; English.

XX The invention relates to isolated human G-protein coupled receptor

XX (GCREC) polypeptides and their biologically active fragments. GCREC and

XX protein is useful in treating a disease or condition associated with an

XX increase or decrease in expression of functional GCREC. The GCREC's are

XX useful in the diagnosis, treatment and prevention of cell proliferative

XX disorders (cancer, leukaemia, melanoma); neurological disorders (stroke,

XX epilepsy, Parkinson's disease, dementia, Alzheimer's disease); autoimmune

XX inflammatory disorder (thyroiditis, haemolytic anaemia, AIDS, multiple

XX sclerosis); cardiovascular disorder (atherosclerosis, angina pectoris),

XX gastrointestinal disorder (ulcer, cholelithiasis, gastroenteritis),

XX metabolic disorders (diabetes); viral infections (herpes virus) and in

XX the assessment of the effects of exogenous compounds on the expression

XX of the nucleic acid and amino acid sequences. The present sequence is

XX human GCREC-3 cDNA.

XX Sequence 1542 BP; 428 A; 327 C; 315 G; 472 T; 0 other;

XX Query Match 38.4%; Score 592.4; DB 24; Length 1542;

XX Best Local Similarity 75.3%; Pred. No. 1.2e-139;

XX Matches 764; Conservative 0; Mismatches 246; Indels 4; Gaps 2;

XX 39 GCAGAAATGCGACAGAAATTTATCTTGTGAGAAATTTGGTGGCAACAGAGCGCTATCTTGAATA 98

XX 205 GGATCATGCGATGGAATGCAACTTGCAAAACACTGGCTGGCAGCAGAGCGCTCCCTGGAAA 264

DB

QY 99 AGTACTACCTCTCTGCAATTTTATGCAATCGAGTTTCATTTTGGACTGCTTGGGAATGCA 158

DB 265 AGTACTACCTTTCCATTTTATGGGATTCGTTGGGAGTCTTGGAAATACCA 324

QY 159 CTGTGGTGTGGGCTACTCTTCTGCAATGAGAACTGGAACAGCAAGTCTATCTTT 218

DB 325 TTGTGTTTACGGGTACATCTTCTCTGAAAGAACTGGAACAGCAAGTCTATCTCT 384

QY 219 TTAACCTTTTCCATCTCTGCAATTTTCTCTGCAACCTTCCCATCTCTGATAAAGATT 278

DB 385 TTAACCTCTCTCTCTGCAATTTTCTCTGCAACCTTCCCATCTCTGATAAAGATT 444

QY 279 ATGCCAATGATAAGGGGACCTTATGAGATGTTCTCTGTAATAAGCAACGATATGCTTC 338

DB 445 ATGCCAATGGAACCTGGATATATGAGACGCTCTCTGCATAAGCAACGATATGCTTC 504

QY 339 ACACCAACCTCTACACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398

DB 505 ATGCCAACCTCTATACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 564

QY 399 TCATGAAGTACCTTTCCGAGAACACTTTCTACAAAGAAAGAAATTTGCCATTTTAACT 458

DB 565 TAAATGATATCTCTCCGAGAACACTTTCTGCAAGAAAGAAAGAAAGTCTCTATTTAACT 624

QY 459 CGCTGGCTGTCTGGGCTTATGAGCTTTAGAGTTCTACCCATCTCTCTCTCTCTCTCT 518

DB 625 CCTTGGCCATTTGGGTTTATGAGCTTTAGAGTTTACTACCCATCTCTCTCTCTCTCT 684

QY 519 CTGTCCCAAAAGAGAGGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 578

DB 685 CTGTTATATGATGCAATGGCACCACTCTGTAATGATTTTGAAGTTCTGGAGACCCCACT 744

QY 579 ACAATCTCAATTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 638

DB 745 ACAACCTCAATTTACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 804

QY 639 TGTGCTTCTTCTACACAGATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 698

DB 805 TGTGTTTCTTTTATACAGATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 864

QY 699 CTGCTCTGCACTGGACAAACCCCAACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 758

DB 865 CTGCTCTGCTCTTGAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 924

QY 759 TACTCTTCAACCTCTATATATCATGCGCAATTTGAGGATCGCTCTCTCTCTCTCTCT 818

DB 925 TGTCTTTTACACCTTATACGCTATGCGGAATGTTGAGGATCGCTCTCTCTCTCTCT 984

QY 819 G---GCCACAAGGATGTACACAGAAGCCCAATCTATATACACATGACACGCGCTC 875

DB 985 GGAAGCAGTATCAGTGCATCAGGTCGTCATCACTCTCTCTCTCTCTCTCTCTCTCT 1044

QY 876 TGGCTCTTCTGAACAGTGCATCAATCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 935

DB 1045 TGGCTCTTCTGAACAGTGCATCAATCCCATCTCTCTCTCTCTCTCTCTCTCTCTCT 1104

QY 936 GAGAGATGCTGATTAAGTTCAGCAATCTCTCAAGTCCCTCTCTCTCTCTCTCTCTCT 995

DB 1105 GGGACATGCTGATGAATCACTGAGACAACTCTCAATCCCTCTCTCTCTCTCTCTCT 1164

QY 996 GAGCTCTGGATGCGAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1048

DB 1165 GGCTCATGNACT 1218

RESULT 6

AAT71900

ID AAT71900 standard; cDNA; 1996 BP.

XX

AC AAT71900;

XX

DT 11-SEP-1997 (first entry)

XX Human purinergic receptor P2U2 cDNA.
 DE P2U2 receptor; purinergic receptor; diagnosis; therapy; ss.
 KW Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FT CDS 625..1629
 FT /*tag= a
 XX
 XX W09720045-A2.
 PN
 XX
 XX 05-JUN-1997.
 PD
 XX
 XX 08-NOV-1996; 96WO-US18175.
 XX
 XX 15-NOV-1995; 95US-0559524.
 PR 15-NOV-1995; 95US-0006782.
 XX
 XX (CORT-) COR THERAPEUTICS INC.
 PA
 XX Conley PB, Jantzen H;
 PI
 XX WPI: 1997-310601/28.
 DR P-PSDB; AAW19854.
 XX
 XX New isolated purinergic receptor sub-type - used to develop
 PT products for diagnosis and therapy, e.g. for screening for agonists
 PT and antagonists which can modulate activation
 PT
 XX Claim 3; Fig 1A-C; 36pp; English.
 PS
 XX
 CC A cDNA clone (AA171900) codes for a novel human purinergic receptor
 CC subtype, designated P2U2 receptor (AAW19854), that is abundantly
 CC expressed in kidney and in many cell lines of megakaryocytic or
 CC erythroleukemic origin and which is activated by ATP, UDP, UTP and
 CC UDP. The clone was obtcd. by amplifying DAMI (ATCC CRL 9792) cell
 CC cDNA using primers (see also AAT72104-05) based on transmembrane
 CC regions of mouse P2u and chicken P2Y1 receptors, and use of the PCR
 CC product to screen the DAMI cDNA library to isolate the full-length
 CC clone. P2U2 nucleic acids can be used in the recombinant prodn. of
 CC P2U2 receptor polypeptides and as probes.
 XX
 SQ Sequence 1996 BP; 513 A; 454 C; 381 G; 647 T; 1 other;

Query Match 38.2%; Score 589.2; DB 18; Length 1996;
 Best Local Similarity 75.1%; Pred. No. 8.9e-139;
 Matches 762; Conservative 0; Mismatches 248; Indels 4; Gaps 2;

QY 39 GCAGATGCGCAGAGATTTATCTTGTGAGAAATGGTTGGCAACAGAGGCTATCTTGATA 98
 DB 632 GGATCATGCGATGGAATGCAATTCGAAATGCTGCTGAGAACTGGCTGGCAGAGAGGCTGCCCTGGAAA 691
 QY 99 AGTACTACTCTCTCGATTTATGCAATCGAGTTCATTTTGGAGTGTCTTGGAGTGTCTTGGAGATGTCA 158
 DB 692 AGTACTACTCTCTCGATTTATGGAATGAGTGTCTTGGAGTGTCTTGGAGTGTCTTGGAGATGTCA 751
 QY 159 CTGTGGTGTGGCTTACCTCTCTCTGATGAAAGAACTGGAACAGAGCAATGCTATCTT 218
 DB 752 TTGTTGTTTACGGTACATCTCTCTGAGAACTGGAACAGAGCAATGATTTATTTCTCT 811
 QY 219 TTAACCTTTCATCTCTGACCTTGTCTTGTGACCTTCCCATCTCTGATAAAGAGTT 278
 DB 812 TTAACCTTCTCTCTGACCTTGTCTTGTGACCTTCCCATCTCTGATAAAGAGTT 871
 QY 279 ATGCAATGATAAAGGGACCTATGAGATGTCTCTGATTAAGCAACCGATATGTGCTTC 338
 DB 872 ATGCAATGGAACCTGGATATATGAGACGTGCTCTGATTAAGCAACCGATATGTGCTTC 931
 QY 339 ACACCAACCTCTACACGAGATCTCTCTCTCACTTTTATAGCATGACCATGATATCTGC 398
 DB 932 ATGCAACCTCTATACGAGATCTCTCTCTCACTTTTATAGCATGATGATCTTGA 991

QY 399 TCATGAAGTACCCCTTTCCGAGAACACACTTCTACAAAAGAGAAATTTGCCATTTAATCT 458
 DB 992 TAATTAAGTATCTCTTTCCGAGAACACCTTCTGCAAAAGAGAGTTTCTGCTATTTAATCT 1051
 QY 459 CGCTGGCTGTCGCGCTTAGTGACCTTAGAGTTCTAGAGTTCTACCCATGCTCACTTTCAATCAAT 518
 DB 1052 CTTTGGCATTGTTGGTTTATAGTAACCTTAGAGTTTACTACCCATGCTCTCCCTTATAAATC 1111
 QY 519 CTGTCCCAAGAGAGGGCAGTAACATGCATCGACTATGCAAGTTCTGGAACCCCTGAAC 578
 DB 1112 CTGTTATAAATGACAATGSCACCACTCTAATGATTTTGAAGTTCTGAGACCCCACT 1171
 QY 579 ACAATCTCATTTACAGCCTCTGCTGAGCTTTGTTGGGTTCTTAAATCTCTCTCTGTA 638
 DB 1172 ACAACCTCATTTACAGCATGTCTTAACACTGTTGGGTTCTTATCTCTCTTTTGTGA 1231
 QY 639 TGTGCTTCTTCTACTACAGATGTTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCA 698
 DB 1232 TGTGTTCTTTTATACAGATGCTCTCTTCTTAAAGCAGAGGAATAGGCGAGTTGCTA 1291
 QY 699 CTGCCCTGCCACTGGACAAACCCCAAGCCTGGTGGTCTCTGGCGTTGTGATCTTCTCTA 758
 DB 1292 CTGCTCTGCCCTTTGAAAGCCTCTCAACTTGGTCATCGCAGTGGTAATCTTCTCTG 1351
 QY 759 TACTCTTCACACCCCTATCATATATCATGCGCAATTTGAGGATCGCCTCAGGCTGTGATAGT 818
 DB 1352 TGTCTTTTACACCTTATCATGCTCATGGGAATGTGAGGATCGCTTACGCTTGGGAGTT 1411
 QY 819 G---GCCACAAGATGTACAGAGGCGCATCAAAATCTATATACACACTGACAGCGGCTC 875
 DB 1412 GGAAGCAGTATCATGCTGCACTCAGGTCGTCATCACTCTTTTACATTTGACAGGCGCT 1471
 QY 876 TGGCCTTTCTGAACAGTGCATCAATGCCATCTTCTTCTCTCTCATGCGGAGACCATTACA 935
 DB 1472 TGGGCTTTCTGAACAGTGTCTCAACCCCTGTCTTCTTCTTCTTTTGGGAGATCACTTCA 1531
 QY 936 GAGAGATGCTGATTAGTAAGTTTACAGAAATCTTCAAGTCCCTTACATCTCTCAGGACAT 995
 DB 1532 GGGACATGCTGATGAATCAACTGAGACACACATCAATCCCTTACATCTCTTAGCAGAT 1591
 QY 996 GAGTGTGTGGATGAGGCTTCTTCACTACGCCAATAA-TGAGACACTTGTATAAAGACAG 1048
 DB 1592 GGGCTCATGAACCTCTTCTTCACTTATCAGAGAAAAGTCAAGGCTTGTGAAACAG 1645

RESULT 7
 AAT75146
 ID AAT75146 standard; cDNA; 1428 BP.
 XX AC AAT75146;
 XX XX
 XX XX
 XX 07-OCT-1997 (first entry)
 XX Human ATP receptor cDNA.
 DE ATP receptor; G-protein coupled receptor; agonist; antagonist; ss.
 KW Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT CDS 92..1096
 FT /*tag= a
 FT /transl_except= (pos:725..727, aa:Ser)
 FT /transl_except= (pos:764..766, aa:Ser)
 FT /transl_except= (pos:820..822, Xaa)
 FT /notes= "Xaa = unknown"
 FT complement (92..100)
 FT /*tag= b
 FT /note= "binding site for primer used to amplify
 FT cDNA for baculovirus expression"
 FT primer_bind complement (92..109)
 FT /*tag= c

FT /note= "binding site for primers used to amplify
FT CDNA for bacterial or COS expression"
FT primer_bind 1076..1095
FT /*tag= d
FT /note= "binding site for primer used to amplify
FT CDNA for COS expression"
FT primer_bind 1079..1096
FT /*tag= e
FT /note= "binding site for primer used to amplify
FT CDNA for bacterial expression"
FT primer_bind 1085..1096
FT /*tag= f
FT /note= "binding site for primer used to amplify
FT CDNA for baculovirus expression"
FT XX
PN WO9724929-A1.
XX
PD 17-JUL-1997.
XX
XX 11-JAN-1996; 96WO-US00392.
XX
XX 11-JAN-1996; 96WO-US00392.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI LI Y;
XX
DR WPI; 1997-372505/34.
DR P-PSDB; AAW22732.
XX
PT Isolated human ATP receptor - agonists and antagonists of which are
PT useful in treatment of, e.g. asthma, hypertension, arterial
PT thrombosis and psychotic and neurological disorders
XX
PS Claim 7; Fig 1A-C; 53pp; English.
XX
CC A CDNA clone (AA75146) codes for human ATP receptor (AAW22732), a
CC polypeptide structurally related to the G protein-coupled receptor
CC family. It was discovered in a human placenta cDNA library.
CC cDNA encoding the mature receptor, deposited as ATCC 97333, can
CC be expressed in bacterial (e.g. E. coli), mammalian (e.g. COS) or
CC insect (e.g. Sf9) host cells and used to screen for agonists and
CC antagonists useful in the treatment of a variety of disorders.
CC It can also be used to identify a mutation in an ATP receptor gene
CC and thus to diagnose diseases, or susceptibility to diseases,
CC related to ATP receptor underexpression.
XX
SQ Sequence 1428 BP; 394 A; 308 C; 290 G; 435 T; 1 other;

Query Match 38.1%; Score 587.2; DB 18; Length 1428;
Best Local Similarity 75.0%; Pred. No. 2.5e-138;
Matches 760; Conservative 1; Mismatches 249; Indels 4; Gaps 2;

QY 39 GCAGAAATGGCAGAGAAATTTATCTGTGAGAAATGGTTGGCAACAGAGAGGCTATCTTGAATA 98
DB 99 GGATCATGGCATGGAATGCAACTTGGCAAACTGGCTGGCGAGAGAGGCTGCCCTGGAAA 158
QY 99 AGTACTACCTCTCGCAATTTATGCAATCGAGTTCAATTTTGGAGCTGCTGGAAATGCTCA 158
DB 159 AGTACTACCTTTCCATTTTATGGGATTTGAGTTGCTGTGGAGTCTTTGGAAATACCA 218
QY 159 CTGTGGTGTTCGGTACCTCTCTCGCATGAAGAACTGGAACACAGCAATGTCTATCTTT 218
DB 219 TTGTGTGTGTACGGTATCTCTCTCTGAAGAACTGGAACAGCAATGATTAATTAATCTCT 278
QY 219 TTAACCTTTTCCATCTCGACTTTTCTCTCTGTCACCTTCCCATCTCTGATAAAGAGTT 278
DB 279 TTAACCTCTCTCTGACTTACCTTTTCTGTGACCTTCCCATCTCTGATAAAGAGTT 338
QY 279 ATGCCAATGATAAGGGACCTATGAGATGTTCTCTGTATAAGCAACCGATATGCTGCTC 338
DB 339 ATGCCAATGGAACCTGGATATATGAGACGCTCTCTGCAATGCAACCGATATGCTGCTTC 398

QY 339 ACACCAACCTCTACACAGAGCATC 398
DB ATGCAACCTCTATACAGCATCT 458
QY 399 TCATGAAGTACCTCTTCCGAGAACACTTCTCTACAAAAGAGGAATTTGCCATTTTAACTCT 458
DB 459 TAATTAAGTATCTCTTCCGAGAACACTTCTCTGAAAAGAGAGAGTGTCTATTTTAACTCT 518
QY 459 CGCTGGCTGTCTGGGCTTGTAGTACCTTTAGAGTTTCTACCCATGCTCTCTCTCTCTCTCT 518
DB 519 CTTGGCCATGTGGTTTGTAGTAACTTTAGAGTTTCTACCCATGCTCTCTCTCTCTCTCTCT 578
QY 519 CTGTCCCAAGAGAGAGGAGGAGTAACTGATGCACTATGCAAGTTCTGGAACCCCTGAAC 578
DB 579 CTGTTATAACTGACATGGCACCCTGTAAATGATTTTGAAGTTCTGGAGAGCCCACT 638
QY 579 ACAATCTCATTTACAGCTCTGCTGACTTTTGTGGGCTTCTCTCTCTCTCTCTCTCTCTCT 638
DB 639 ACAACCTCATTTACAGCATGTGCTAACACTGTTGGGTTCTTATTTCTCTCTCTCTCTCT 698
QY 639 TGTCT 698
DB 699 TGTCT 758
QY 699 CTGCTCTGCTCTGCAACCT 758
DB 759 TACT 818
QY 819 TGCYTTTACACCTCTATCATGCGCAATTTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCT 878
DB 879 GGAAGCAGTATCATGTCAGTCT 938
QY 876 TGGCT 935
DB 939 TGGCT 998
QY 936 GAGAGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 995
DB 999 GGGAGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1058
QY 996 GAGCTGTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1048
DB 1059 GGGCTCATGAACT 1112

RESULT 8
AAC81122
ID AAC81122 standard; cDNA; 1385 BP.
XX
AC AAC81122;
XX
DT 14-FEB-2001 (first entry)
XX
DE Human secreted protein gene 37 SEQ ID NO:47.
XX
KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; vulnery; gene therapy; autoimmune disease;
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW cerebrovascular disorder; nervous system disorder; infection; skin aging;
KW ocular disorder; wound healing; food additive; preservative; ss.
XX
OS Homo sapiens.
XX
PN WO200061628-A1.
XX
PD 19-OCT-2000.
XX

PD	25-JUL-2002.	
XX	17-JAN-2002; 2002WO-JP00270.	
PF	18-JAN-2001; 2001JP-0010714.	
XX	30-MAR-2001; 2001JP-0102484.	
PR	(TAKE) TAKEDA CHEM IND LTD.	
XX	Miwa M, Ito T, Shintani Y, Miyajima N;	
PI	WPI; 2002-566800/60.	
DR	P-PSDB; AAO15399.	
XX	Human kidney-originated G protein-coupled receptor protein TGR30 and	
PT	encoded DNA, for developing drugs to treat central nervous diseases,	
PT	endocrine diseases, metabolic diseases and cancer, including gene	
PT	therapy	
XX	Claim 6; Page 90-91; 98pp; Japanese.	
PS	The invention comprises the amino acid and coding sequence of a human G	
XX	protein-coupled receptor. The DNA and protein sequences of the invention	
CC	are useful for developing drugs to prevent or treat (gene therapy):	
CC	central nervous system diseases; endocrine diseases; metabolic diseases;	
CC	cancer; respiratory diseases; digestive diseases; immune diseases;	
CC	inflammations; infections; and circulatory diseases. The present DNA	
CC	sequence encodes the human G protein-coupled receptor of the invention.	
XX	Sequence 1011 BP; 257 A; 263 C; 188 G; 303 T; 0 other;	
SQ		
Query Match 8.2%; Score 126.6; DB 24; Length 1011;		
Best Local Similarity 49.9%; Pred. No. 6.7e-22;		
Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;		
QY	60 CTTGTGAGAAATGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCAATTTT 119	
DB		
QY	59 CTTTGGAAATGGCACTGATGAAACATCCCACTCAAGATGCACCTACCTCCCTGTTATTT 118	
DB		
QY	120 ATGCAATCGAGTTTCATTTTGGAGCTCTTGGGAATGTCACTGTGTGGTTCGGCTACCTCT 179	
DB		
QY	119 ATGGCAATATCTCTCTCGTGGATTTCCAGGCAATGCAGTAGTATGCCACTTACATTT 178	
DB		
QY	180 TCTGCAATGAAGAACTGGACAGCAAGATGCTATCTCTTTTAACTTCCATCTCTGACT 239	
DB		
QY	179 TCAAAATGAGACCTTGGAGAGAGCAGCACCATTATGCTGAACCTGGCCCTGCACAGATC 238	
DB		
QY	240 TTGCTTTCTGTGCACCTTCCCATCTGATAAGAGTTATGCCAAT---GATAAGGGGA 296	
DB		
QY	239 TGCTGTATCTGACCACTCCCTCTCTGATTCACCTATGCTGCTGCTGCTGCTGCTGCTG 298	
DB		
QY	297 CCTATGGAGATGTTCTCTGTATAAGCAACCGATATGCTTCCACCAACCTCTACACCA 356	
DB		
QY	299 TCTTTGGAGATTTTCATGTTGTAAGTTTATCCGCTTCACTTCCATTTCAACCTGTATAGCA 358	
DB		
QY	357 GCATCTCTTCTTCATTTTATAGCATGAGCGGATATCTGCTCATGAAGTACCCCTTCC 416	
DB		
QY	359 GCATCTCTTCTTCATTTTATAGCATGAGCGGATATCTGCTCATGAAGTACCCCTTCC 418	
DB		
QY	417 GAGAACACATTTCTACAAAAGAGAAATTTGCCATTTTAACTCGCTGCTGCTGCTGCTG 476	
DB		
QY	419 GCTGCTTTTCCATTCACAAACTCGATGTCAGTTGTAGCTGTACCTGTGCTGTGCTGATCA 478	
DB		
QY	477 TAGTGACCTTTAGAGTTCTACCCATGCTCACTTTTCAATTTGTCCCAAGAGAGG 536	
DB		
QY	479 TTTCACTGGTAGTGTCATTTCCGATGACCTTCTTGATCATCAACCAACAGGACCAACA 538	
DB		
QY	537 GCAGTAAGTGCATGCGATATGGAAGTTCTGGAAACCTTGAACACATCTCAATTACAGCC 596	
DB		
QY	539 GATCAGCTCTCTCGACCTCACCAGTTCGG-----ATGAACCTCAATCACTAATTAAGTGT 592	
DB		
QY	597 TCTGCCCTGACTTTGTTGGGCTTCCCTAATCTCTCTCTGTGATGCTGCTTCTTCTACTACA 656	
DB		

Db	593	ACAACTGATTTTGACTGCAACTACTTTTCTGCTCCCTTGTGTGATGAGTACACTTTGCT	652
QY	657	AGATGCTAGTCTCTTTAAAGAGGAGGAGCAGCAGCAAGCAACTGCCCTGCCACTGGACA	716
Db	653	ATACCAGATTTATCCACTCTGACCTTGCACCTTGCACCTTGCCTTTAAGCAGA	712
QY	717	AACCCCAAGCCTGTTGGTGGCGGTTGTGATCTTCTTACTTCTTACACCTTATC	776
Db	713	AAGCAGCAAGGCTTAACCAATCTCTGCTTACTGCTTGTGATTTTACGTATGTTTTTACCTTCC	772
QY	777	ATATCATGCGCAATTTGAGATGCGCTCACGCTG	811
Db	773	ATATCTTGGGTCATTCGGATCGAATCTCGCCTG	807
RESULT 10			
AAS07948			
ID	AAS07948	standard; cDNA; 1014 BP.	
XX	AAS07948;		
XX	23-OCT-2001 (first entry)		
DE	Human cDNA encoding G-protein coupled receptor, hrUP21.		
XX	Human; G-protein coupled receptor; GPCR; hrUP21; agonist;		
KW	inverse agonist; lung cancer; ss.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	CDS	1..1014	
FT	FT	/tag= a	
FT	FT	/product= "hrUP21"	
XX	WO200136471-A2.		
XX	25-MAY-2001.		
XX	16-NOV-2000; 2000WO-US31509.		
PR	17-NOV-1999; 99US-0166088.		
PR	17-NOV-1999; 99US-0166099.		
PR	23-DEC-1999; 99US-0171900.		
PR	23-DEC-1999; 99US-0171901.		
PR	23-DEC-1999; 99US-0171902.		
PR	11-FEB-2000; 2000US-0181749.		
PR	14-MAR-2000; 2000US-0189258.		
PR	14-MAR-2000; 2000US-0189259.		
PR	10-APR-2000; 2000US-0195898.		
PR	10-APR-2000; 2000US-0195899.		
PR	10-APR-2000; 2000US-0196078.		
PR	28-APR-2000; 2000US-0200419.		
PR	12-MAY-2000; 2000US-0203630.		
PR	12-JUN-2000; 2000US-0210741.		
PR	21-AUG-2000; 2000US-0226760.		
PR	26-SEP-2000; 2000US-0235418.		
PR	26-SEP-2000; 2000US-0235779.		
PR	20-OCT-2000; 2000US-0242332.		
PR	20-OCT-2000; 2000US-0242343.		
XX	(AREN-) ARENA PHARM INC.		
PA	Chen R, Dang HT, Lowitz KP;		
XX	WPI; 2001-355616/37.		
DR	P-PSDB; AAU04375.		
XX	Endogenous and non-endogenous versions of human G-protein coupled		
PT	receptors for direct identification of candidate compounds as agonists,		
PT	inverse agonists or partial agonists for use as therapeutic agents -		

RESULT 11	
ABN85630	
ID	ABN85630 standard; DNA; 1014 BP.
XX	
AC	ABN85630;
XX	
DT	18-SEP-2002 (first entry)
XX	
DE	Human P2Y-like receptor variant encoding gene SEQ ID NO 3.
XX	
KW	Human; P2Y-like receptor; HIPHUM 0000037; immunity; inflammation;
KW	cancer; Crohn's disease; irritable bowel syndrome; rheumatoid arthritis;
KW	immunomodulator; anti-inflammatory; cytostatic; antiasthmatic;
KW	gastrointestinal; anti-ulcer; antirheumatic; antiarthritic; virucide;
KW	antibacterial; immunosuppressive; dermatological; nephrotropic;
KW	antiallergic; analgesic; receptor; gene; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	1..1014
FT	/*tag= a
FT	/product= "P2Y-like receptor variant"
XX	
PN	GB2369364-A.
XX	
PD	29-MAY-2002.
XX	
PF	31-AUG-2001; 2001GB-0021215.
XX	
PR	01-SEP-2000; 2000GB-0021524.
PR	06-SEP-2000; 2000GB-0021894.
PR	25-SEP-2000; 2000GB-0023444.
XX	
PA	(GLAX) GLAXO GROUP LTD.
XX	
PI	Foord SM, Ignar DM;
XX	
DR	WPI; 2002-511268/55.
XX	
DR	P-PSDB; ABR83819.
XX	
PT	An isolated p2y-like receptor polypeptide (HIPHUM 0000037) which can be
PT	used for the identification of agonists and antagonists which may be
PT	used to treat an immune or inflammatory disease -
XX	
PS	Claim 5; Page 28-29; 35pp; English.
XX	
CC	The invention relates to an isolated p2y-like receptor polypeptide
CC	(ABR83818-ABR83819) which is also referred to in the specification as
CC	HIPHUM 0000037. An effective amount of a substance (agonist or
CC	antagonist) which modulates p2y receptor activity is useful to treat a
CC	subject having a disorder that is responsive to p2y-like receptor
CC	modulation. The disorder is a disease of immunity or inflammation. The
CC	substance may also be used to manufacture a medicine for the treatment of
CC	prophylaxis of a disorder that is responsive to stimulation or modulation
CC	of p2y-like receptor activity. Disorders which may be treated include
CC	colon cancers, asthma, COPD, Crohn's disease, irritable bowel syndrome,
CC	gastroenteritis and colitis, inflammatory bowel syndrome, ulcerative
CC	colitis, rheumatoid arthritis, viral diseases, bacterial infections,
CC	autoimmune diseases, dermatitis, glomerulonephritis allergies, allergic
CC	rinitis, inflammatory pain and general inflammation such as tendonitis,
CC	polymyositis or prostatitis. The invention provides alternative
CC	substances for the treatment of immunological and inflammatory diseases.
CC	The present sequence is that the p2y-like receptor variant encoding gene
CC	of the invention.
XX	
XX	Sequence 1014 BP; 258 A; 263 C; 189 G; 304 T; 0 other;
XX	
Query Match	8.28; Score 126.6; DB 24; Length 1014;
Best Local Similarity	49.99; Pred. No. 6.7e-22;
Match	100.00; Pred. No. 6.7e-22;

60 CTTGTGGAATTGGTTGGCAACACAGGCTATCTTGAATAAGTACTACTCTCTGCAATTTT 119

Db 59 CTTTGGAAATGGCACTGATGAAACATCCCACTCAAGATGCATCTCCTCTTATTT 118
QY 120 ATGCAATCGAGTTCAATTTTGGAGTGTGGAAATGTCAGTGTGGTGGCTACCTCT 179
Db 119 ATGGCAATATCTCTCTGCTGGGATTTCCAGCAATGCAAGTAGTATGATATCCACTTACATTT 178
QY 180 TCTGATGAGAACTGGGACAGCAAGTCTATCTTTTAACTTTCCATCTCTGACT 239
Db 179 TCAAAATGAGACCTTGGGAGGACAGCAACCATATATGCTGAACCTGGCTGCACATC 238
QY 240 TTGCTTTCTGTGACCTTCCATCTCTGATAAAGAGTATGCGCAAT---GATAAGGGGA 296
Db 239 TCGTGTATCTGACAGCTCCCTCTCTGATTCACACTATGCCAGTGGCGGAAACATGGA 298
QY 297 CCTATGGAGATGTTCTCTGTATAAGCAACCGATATGCTTCAACCAACCTTACACCA 356
Db 299 TCTTTGGAGATTTTCATGTTAAGTTTATCCGCTTCAGCTTCCATTTCAACCTGTATAGA 358
QY 357 GCATCTCTCTCTCACTTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCTTTCC 416
Db 359 GCATCTCTCTCTCACTTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCTTTCC 418
QY 417 GAGAACACTTTTACAAAAGAGGAATTTGCCATTTTAACTTCGCTGCTGTCTGGGCT 476
Db 419 GCTGCTTTTCCATTACAAAACCTCGATGTCAGTTGTAGCCTGTCTGTGGTGGATCA 478
QY 477 TAGTGACCTTAGAAGTTTACCACTGCTCACTTTTCATCAATTTCTCCCAAGAAAGAGG 536
Db 479 TTTCACTGTGATGCTGTCATTCCTCGATGACCTTTCTGATCATACCAACAGGACCAACA 538
QY 537 GCAGTAACCTGATCGACTATGAAAGTTCTGGAACCCCTGAACACAACTCTCACTTACAGCC 596
Db 539 GATCAGCTGTCTGACCTCACCACTGCGG-----ATGAACCTCAATCTATTAAGTGT 592
QY 597 TCTGCTGATGTTGGGCTCTCAATTTCTCTCTCTGATGCTCTCTCTCTCTCTCTCTCT 656
Db 593 ACAACCTGATTTTGAAGTGAACCTTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 652
QY 657 AGATGTAGTCTTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 716
Db 653 ATACAGGATATTCACACTCTGACCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 712
QY 717 AACCCCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
Db 713 AAGCAGGAGGCTAACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 772
QY 777 ATATCATGCCAATTTGAGGATGCTCTCAGGCTG 811
Db 773 ATATCTTGGAGGCTCATTCGGATGCAATCTCGCCTG 807

RESULT 12
ABK11381
ID ABK11381 standard; DNA; 1014 BP.
XX
AC ABK11381;
XX
XX
DT 05-JUN-2002 (first entry)
XX
DE Human DNA encoding P2Y1-like G protein-coupled receptor.
XX
KW Human; ds; gene; P2Y1-like G protein-coupled receptor; GPCR;
KW infection; pain; cancer; anorexia; bulimia; asthma; hypotension;
KW central nervous system disease; acute heart failure; hypertension;
KW urinary retention; osteoporosis; diabetes; angina pectoris;
KW myocardial infarction; ulcer; inflammation; allergy; multiple sclerosis;
KW benign prostatic hypertrophy; psychosis; neurological disorder;
KW dyskinesia; HIV; human immunodeficiency virus infection; CNS disorder;
KW Parkinson's disease; anxiety; schizophrenia; manic depression; delirium;
KW dementia; severe mental retardation; Huntington's disease;
KW Tourette's syndrome.

OS Homo sapiens.
XX Key Location/Qualifiers
FH 1..1014
FT CDS /*tag= a "P2Y1-like GPCR"
FT /product= "P2Y1-like GPCR"
XX
PN W0200214511-A2.
XX
XX 21-FEB-2002.
XX
XX 10-AUG-2001; 2001WO-BP09243.
XX
XX 14-AUG-2000; 2000US-224989P.
XX
XX (FARB) BAYER AG.
XX
XX Ramakrishnan S;
XX
XX WPI; 2002-257607/30.
XX P-PSDB; AAU77600.
XX
XX Novel human P2Y1-like G protein-coupled receptor polypeptide which can
XX be regulated for treating infection, pain, cancer, diabetes, anorexia,
XX asthma, hypertension, neurological disorder and dyskinesia
XX
XX Claim 1; Fig 5; 118pp; English.
XX
XX The invention relates to a purified human P2Y1-like G protein-coupled
XX receptor (GPCR) polypeptide and the nucleic acids encoding it
XX (including 5' and 3' sequences, promoters, fragments, variants, or a
XX sequence encoding a protein at least 50% identical to the GPCR).
XX Also included are an expression vector comprising the nucleic acid,
XX a host cell containing the vector and the identification of modulators of
XX the GPCR especially those that reduce the activity of the GPCR.
XX The nucleic acid is useful for detecting a polynucleotide encoding
XX the GPCR in a biological sample. The GPCR and nucleic acid are useful for
XX screening for agents which decrease the activity of the GPCR and
XX for modulators of the GPCR. The modulator or agent useful for modulating
XX the activity of P2Y1-like G protein-coupled receptor in a disease such as
XX bacterial, fungal, protozoan, and viral infection, pain, cancer,
XX anorexia, bulimia, asthma, central nervous system (CNS) disease, acute
XX heart failure, hypotension, hypertension, myocardial infarction, ulcer,
XX osteoporosis, diabetes, angina pectoris, benign prostatic hypertrophy,
XX inflammation, allergy, multiple sclerosis, dyskinesias, HIV virus infection
XX (human immunodeficiency virus), CNS disorders such as Parkinson's
XX disease, anxiety, schizophrenia, manic depression, delirium, dementia,
XX severe mental retardation, Huntington's disease and Tourette's syndrome.
XX The present sequence encodes the P2Y1-like GPCR of the invention.
XX
XX Sequence 1014 BP; 258 A; 263 C; 189 G; 304 T; 0 other;
XX
XX Query Match 8.2%; Score 126.6; DB 24; Length 1014;
XX Best Local Similarity 49.9%; Pred. No. 6.7e-22;
XX Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;
XX
QY 60 CTTGTGAGAAATGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTCTCTCT 119
Db 59 CTTTGGAAATGGCACTGATGAAACATCCCACTCAAGATGCATCTCCTCTTATTT 118
QY 120 ATGCAATCGAGTTCAATTTTGGAGTGTGGAAATGTCAGTGTGGTGGCTACCTCT 179
Db 119 ATGGCAATATCTCTCTGCTGGGATTTCCAGCAATGCAAGTAGTATGATATCCACTT 178
QY 180 TCTGATGAGAACTGGGACAGCAAGTCTATCTTTTAACTTTCCATCTCTGACT 239
Db 179 TCAAAATGAGACCTTGGGAGGACAGCAACCATATATGCTGAACCTGGCTGCACATC 238
QY 240 TTGCTTTCTGTGACCTTCCATCTCTGATAAAGAGTATGCGCAAT---GATAAGGGGA 296
Db 239 TCGTGTATCTGACAGCTCCCTCTCTGATTCACACTATGCCAGTGGCGGAAACATGGA 298

Db 539 GATCAGCCTGCTCGACCTCACCAGTTCGG-----ATGAACCTCAATACTATTAAAGTGGT 592

QY 597 TCTGCCTGACTTTGTTGGCTTCCCTAAATTCCTCTCTGTGATGCTTCTTCTACTACA 656

Db 593 ACAACCTGATTTGACTGCAACTACTTTCTGCTCCCTTGTGTGATAGTACACTTTTGT 652

QY 657 AGATGCTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCACTGCCCTGCCACTGGACA 716

Db 653 ATACCAGGATATCCACACTCTGACCATGAGCTGCAAACTGACAGCTGCCCTTAAGCAGA 712

QY 717 ACCCCCAAGCCTGGTGGTCTGGGGTGTGATCTTCTCTATACTCTTCAACACCTATC 776

Db 713 AAGCAGGAGGCTAACCATCTGCTACTCTCTGCAATTTTACGATGTTTATACCTTCC 772

QY 777 ATATCATGGCAATTTGAGGATCGCTCAGCGCTG 811

Db 773 ATATCTTGAGGGTCATTCGATCGAATCTCGCGTG 807

RESULT 14

ABL56197

ID ABL56197 standard; cDNA; 1288 BP.

XX ABL56197;

AC.

DT

XX 05-JUL-2002 (first entry)

DE Human P2Y1-11 encoding cDNA.

XX Human; P2Y1-11; chromosome 13; G protein-coupled; receptor;

KW gene therapy; thyroid; gene; ss.

XX Homo sapiens.

XX

XX Key Location/Qualifiers

FT 18..1031

FT CDS

FT /*tag= a

FT /product= "P2Y1-like purine receptor"

XX

PN DEL0046970-Al.

XX

XX 11-APR-2002.

XX

XX 22-SEP-2000; 2000DE-1046970.

XX

XX 22-SEP-2000; 2000DE-1046970.

PR

XX (BRUE/) BRUESS H.

PA (BOEN/) BOENISCH H.

XX

XX Bruess M, Boenisch H;

PI

XX WPI; 2002-353329/39.

DR P-PSDB; ABB79438.

XX

XX New human P2Y111 gene, useful for treatment and diagnosis of associated

PT diseases, and related proteins, antibodies and modulators, encodes G

PT protein-coupled receptor

XX

XX Claim 5; Page 3; 5pp; German.

PS

XX The invention relates to the human P2Y111 gene (I), including its 5' and

CC 3' untranslated regions, located on chromosome 13 and encoding a G

CC protein-coupled receptor. (I) and related mRNA, cDNA, protein, antibodies

CC etc., are used for diagnosis and (gene) therapy of diseases that are

CC (indirectly) associated with (I) or its expression products. No diseases

CC are specified but as (I) is expressed only in thyroid tissue, (I) is

CC presumed to be involved in regulation of thyroid function. The present

CC sequence is that of the P2Y111 encoding cDNA.

XX

SQ Sequence 1288 BP; 347 A; 318 C; 241 G; 382 T; 0 other;

Query Match 8.2%; Score 126.6; DB 24; Length 1288;

Best Local Similarity 49.9%; Pred. No. 7.4e-22;

Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;

QY 60 CTTGTGAGATTTGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCAATTT 119

Db 76 CTTTGGAAATTCAGTGTGAAAGAACCCACTCAAGATGCACTACCTCCCTGTTATTT 135

QY 120 ATGCAATCGAGTTTCATTTTGGACTGCTTGGAAATGTCATCTGGTGTTCGGTACTCTCT 179

Db 136 ATGCATTAATCTCTCTGCTGGGATTTCCAGGCAATGTCAGTAGTGATATCCACTTACATTT 195

QY 180 TCTGCATCAAGAACTGAACAGCAGCAATGCTATCTTCTTAACTTTTCACTCTCTGACT 239

Db 196 TCAAAATGAGACCTTGAAGAGCAGCAGCACTCATTTATGCTGAACCTGGCCCTGCACAGATC 255

QY 240 TTGCTTTTCCGTGACACCTTCCCATCTCTGATAAGAGTTTATGCCAAT---GATAAGGGGA 296

Db 256 TGCTGTATCTGACGACGCTCCCTTCTCTGATTTCTACTATGCGAGTGGCGGAAACTGGA 315

QY 297 CCTATGGAGATGTTCTCTGTATAAGCAACCGATATGCTTTCACACCACTCTTACACCA 356

Db 316 TCTTGGAGATTTTCATGTGTAAGTTTATCCGCTTCAGCTTCCATTTCAAGCTGTATAGCA 375

QY 357 GCATCCTCTTCTCCTCATTTTCAATAGCATGACCGATATCTGCTCATGAAGTACCTTTTC 416

Db 376 GCATCCTCTTCTCCTCATTTTCAAGCATTTTCCGCTTCTGCTGATCATTTACCCCAATGA 435

QY 417 GAGAACACTTTTCTACAAAGAAAGAAATTTGCCATTTTAAATCTCGCTGGCTGTCTGGCCT 476

Db 436 GCTGCTTTTCCATTCACAAACTCGATGTCAGTTGTAGCTGTGCTGTGCTGGATCA 495

QY 477 TAGTGACCTTAGAAGTTCTACCCATGCTCACTTTTCATCAATTTCTGTCCTCCAAAGAGAGG 536

Db 496 TTTCACCTGGTGTCTCATTTCCGATCTTCTTGATACATCAACCAACAGGACCAACA 555

QY 537 GCAGTAACCTGCATGACTATCAAGTTTCTGGAACCTCTGAAACCAATCTCATTTACAGCC 596

Db 556 GATCAGCCTGCTCGACCTCACCCTCAGTTCGG-----ATGAACCTCAATACTATTAAAGTGGT 609

QY 597 TCTGCTGACTTTTGTGGCTTCTCTTAATTCCTCTCTGCTGTGATGCTTCTTCTACTACA 656

Db 610 ACAACCTGATTTTGTGACTGCAACTACTTCTGCTCCCTTGGTGTGATGACACTTTTGTCT 669

QY 657 AGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCACTGCCCTGCCACTGGACA 716

Db 670 ATACCACGATTTATCCACACTCTGACCCATGACCTGCAAACTGACAGCTGCCCTTAAGCAGA 729

QY 717 AACCCCAACGCTGCTGCTGCTGGCGGTTGTGATCTTCTCTATACTTTCACACCTATC 776

Db 730 AAGCAGGAGGCTAACCATTTCTGCTACTCTTGCATTTTACGATGTTTATACCTTCC 789

QY 777 ATATCATGGCAATTTGAGGATCGCTCAGCGCTG 811

Db 790 ATATCTTGAGGGTCATTCGATCGAATCTCGCGCTG 824

RESULT 15

AAS08362

ID AAS08362 standard; cDNA; 1729 BP.

XX AAS08362;

AC

XX 26-SEP-2001 (first entry)

DT Human cDNA encoding G-protein coupled receptor, GPCR 39404.

XX

XX Human; G-protein coupled receptor; GPCR; 39404; immunogen; antibody;

KW Good pasture's syndrome; viral hepatitis; Alzheimer's disease; lymphoma;

KW Heymann nephritis; Paget's disease; Crohn's disease; endometriosis;

KW systemic lupus erythematosus; actinic keratosis; myocarditis;

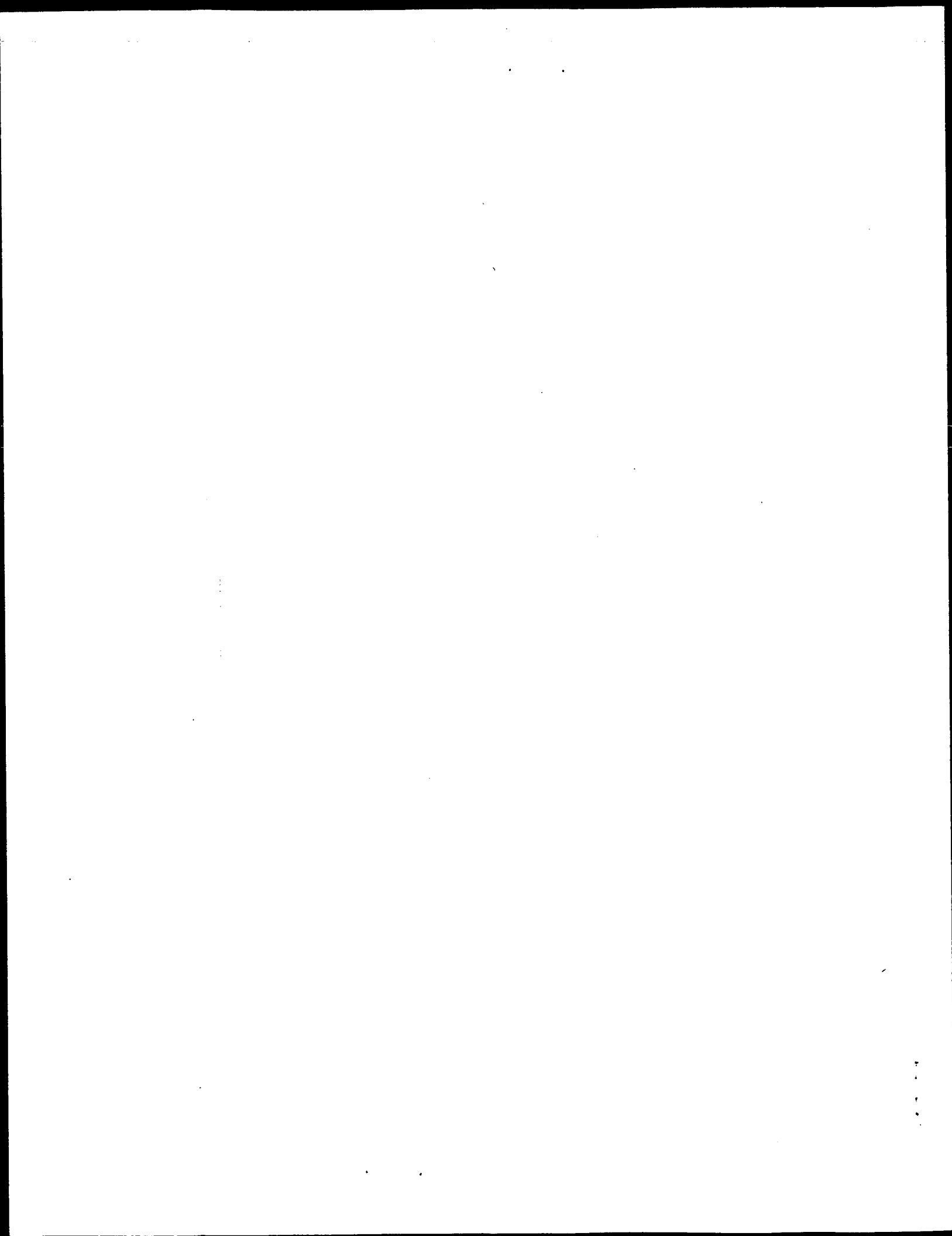
KW Kawasaki syndrome; DiGeorge syndrome; peripheral B-cell neoplasm;

XX inflammations; teratoma; ss.

OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 294..1307
FT /*tag= a
FT /product= "Protein 39404"
XX
XX WO200149847-A2.
XX 12-JUL-2001.
XX
XX 22-DEC-2000; 2000WO-US35309.
XX
XX 30-DEC-1999; 99US-0475790.
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Glucksmann MA, White D;
XX
XX WPI: 2001-432880/46.
XX P-PSDB; AAU04584.
XX
XX Novel isolated 26904, 38911 and 39404 polypeptides which are seven
XX transmembrane proteins belonging to superfamily of G-protein-coupled
XX receptors, useful for treating disorders of spleen, lung, liver, brain
XX and kidney .
XX
XX Claim 2; Fig 1; 164pp; English.
XX
XX The sequence encodes a novel human seven transmembrane domain
XX protein belonging to the G-protein coupled receptor (GPCR) superfamily,
XX protein 39404. The receptor is useful in drug screening assays, to
XX identify compounds that modulate receptor activity and/or interact with
XX the receptor, and for producing antibodies specific for the receptor, its
XX regions or fragments. The receptor is useful for treating/diagnosing a
XX 26904, 38911 and 39404 protein-associated disorder characterised by
XX aberrant expression or activity of the protein, for monitoring
XX therapeutic effect during clinical trials and other treatment, as bait
XX proteins in a two-hybrid or three-hybrid assay, and in pharmacogenomic
XX analysis. The proteins and nucleic acids encoding them are useful for
XX diagnosis and treatment of disorders selected from disorders of the
XX spleen, lung such as Good pasture's syndrome, liver such as viral
XX hepatitis, brain such as Alzheimer's disease, haematopoietic stem cells
XX such as lymphoma, kidney such as Heymann nephritis, bone such as Paget's
XX disease, colon such as Crohn's disease, uterus and endometrium such as
XX endometriosis, T-cell disorders such as systemic lupus erythematosus,
XX diseases of the skin such as actinic keratosis, disorders of the heart
XX such as myocarditis, disorders involving blood vessels such as Kawasaki
XX syndrome, disorders involving the thymus such as DiGeorge syndrome,
XX disorders involving B-cells such as peripheral B-cell neoplasms,
XX disorders of the breast such as inflammations, and disorders involving
XX the testis and epididymis such as teratoma. Numerous examples of
XX each type of disorder are given in the specification.
XX
XX Sequence 1729 BP; 461 A; 403 C; 302 G; 563 T; 0 other;
XX
XX Query Match 8.2%; Score 126.6; DB 22; Length 1729;
XX Best Local Similarity 49.9%; Pred. No. 8.4e-22;
XX Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;
XX
XX QY 60 CTTGTGAGAAATGGTGGCAACAGAGGCTATCTTGAATAGTACTACCTCTCTGCAATTTT 119
XX III III III III III III III III III III III III III III III III
XX DB 352 CTTTGGAAATTCACATGATGAAGAACCTCCACTCAAGATGCACCTACCTCCCTGTATTT 411
XX
XX QY 120 ATGCAATCGAGTTCAATTTTGGAAATCTGGGAATGTCACCTGGGTTCGGCTACCTCT 179
XX III III III III III III III III III III III III III III III III
XX DB 412 ATGCAATATCTTCCTCGGGATTTCCAGGCAATGCAGTAGTGATATCCACTTACATTT 471
XX
XX QY 180 TCTGCATGAAGAATCGAAGACAGCAAGATCTCTATCTTTTAACTTTTCCATCTCTGACT 239
XX III III III III III III III III III III III III III III III III
XX DB 472 TCAAAATGAGACCTTGGAAAGACGACGACCATCATTTATGCTGAACCTGGCCTGCAGATC 531
XX
XX QY 240 TTGCTTCTCTGTCACCCCTTCCCATCTCTGATAAGAGTTATGCCAAT---GATAAGGGGA 296

Db 532 TGCTGTATCTGACCAAGCCTCCCTTCTCTGATTCACCTACTATGCCAGTGGCGAAACTGGA 591
QY 297 CCTATGGAGATGTTCTCTGTATTAAGCAACGATATGCTTACACCAACCTCTTACACCA 356
Db III III III III III III III III III III III III III III III III
592 TCTTTGGAGATTTTCATGTGTAAAGTTTATCCGCTTTCAGCTTCCATTTCAACCTGTATAGCA 651
QY 357 GCATCCTCTTCCCTCACTTTTCATAGCATGGACGGATATCTGCTCATGAAGTACCCCTTCC 416
III III III III III III III III III III III III III III III III
Db 652 GCATCCTCTTCCCTCACTTTTCAGCAATCTTCCGCTACTGTGTGATCATTCACCAATGA 711
QY 417 GAGAACACTTTTCTACAAAAGAAATTTGCCATTTTAAATCTCGCTGGCTGTCTGGGCT 476
III III III III III III III III III III III III III III III III
Db 712 GCTGCTTTTCCATTCACAAAACCTCGATGTGCAGTTGTAGCCTGTGCTGGGTGTGATCA 771
QY 477 TAGTGACCTTAGAAGTTCTACCCATGCTCCTTTTCATCAATTTCTGCTCCCAAGAGAGG 536
III III III III III III III III III III III III III III III III
Db 772 TTTCACTGGTAGCTGTTCATTCGGATGACCTTCTTGATCATCAACCAACGAGCAACA 831
QY 537 GCAGTAACCTGCATCGACTATGCAAGTTCTGGAACCCCTGAACACAATCTCATTTACAGCC 596
III III III III III III III III III III III III III III III III
Db 832 GATCAGCCTGTCTCGACCTCACAGTTCGG-----ATGAACCTCAATACTATTAAGTGT 885
QY 597 TCTGCTGACTTTGTGTGGCTTCTTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACA 656
III III III III III III III III III III III III III III III III
Db 886 ACAACCTGATTTTGACTGCAACTACTTTCTGCTCCCTTGGTGATGTGACACTTTTGTCT 945
QY 657 AGATGCTAGTCTTCTTAAGAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACA 716
III III III III III III III III III III III III III III III III
Db 946 ATACCACGATTATCCACACTCTGACCCATGSGACTGCAAACTGACAGCTGCCCTTAAGCAGA 1005
QY 717 AACCCCAACGCTGTGTGGTGGCTGTGCTGGGGTGTGATCTTCTCTACTCTTTCACACCTATC 776
III III III III III III III III III III III III III III III III
Db 1006 AAGCAGGAAGGCTAACCACTTCTGCTACTCTTCTGCAATTTTACGTATGTTTTTACCCCTCC 1065
QY 777 ATATCATGCGCAATTTGAGGATCGCCTCACGCGCTG 811
III III III III III III III III III III III III III III III III
Db 1066 ATATCTTGAGGCTCATTCGGATCGAATCTCGCCTG 1100

Search completed: April 2, 2003, 13:08:37
Job time : 399 secs



Db 632 GGATCATGGCATGGAATGCAACTTGCAGAAACCTGGCTGGCAGCAGAGGCTGCCCTGGAA 691
QY 99 AGTACTACCTCTCTGCATTTATGCAATCGAGTTTCATTTTGGACTGCTTGGAAATGCA 158
Db 692 AGTACTACCTTTCCATTTTATGGGATGAGTTCGTTGGGAGTCCCTTGGAAATACCA 751
QY 159 CTGTGTGTTGGCTACCTCTCTCTGATGATGAAGAACTGGAGCAAGCAGCATGCTATCTT 218
Db 752 TTGTTGTTTACGGCTACATCTCTCTCTGAAGAACTGGAGCAAGCAGCATGCTATCTT 811
QY 219 TTAACCTTTCCATCTCTGACTTTGCTTCTGTCACCTTCCCATCTCCATCTGATAAAGAGTT 278
Db 812 TTAACCTCTCTGCTCTGACTTACCTTTTCTGTCACCTTCCCATCTCCATCTGATAAAGAGTT 871
QY 279 ATGCCAATGATAAGGGACCTATGGAGATGTTCTCTGATAAGCAACCGATATGCTTC 338
Db 872 ATGCCAATGAACTGGATATATGGAGACGCTCTCTGATAAGCAACCGATATGCTTC 931
QY 339 ACACCAACCTCTACACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398
Db 932 ATGCCAACCTCTATACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 991
QY 399 TCATGAAGTACCTTTCCGAGAACACTTTCTTACAAAGAAAGAAATTTGCCATTTAACT 458
Db 992 TAATTAAGTATCTTTCCGAGAACACTTTCTGCAAAAGAAAGAGTTTGTCTATTTAACT 1051
QY 459 CGTGGCTGTCTGGGCTTTAGTACCTTAGAGTTCTACCCATCTGCTACCTTTTCATCAAT 518
Db 1052 CTTGGCCATTTGGGTTTGTAGTAACTTTAGTAACTTTAGTAACTTTAGTAACTTTAGTAACT 1111

APPLICANT: Jantzen, Hans-Michael
TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,707
FILING DATE: 15-NOV-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 044481-5010-01-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1996 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 625..1626
US-08-749-707-1

Query Match 38.2%; Score 589.2; DB 3; Length 1996;
Best Local Similarity 75.1%; Pred. No. 3.1e-156;
Matches 762; Conservative 0; Mismatches 248; Indels 4; Gaps 2;

QY 39 GCAGAAATGGCAGCAAGATTTATCTTGTGAGAAATTTGTTGGCAACAGAGGCTATCTTGAATA 98
Db 632 GGATCATGGCATGGAATGCAACTTGCAGAAACCTGGCTGGCAGCAGAGGCTGCCCTGGAA 691
QY 99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTTCATTTTGGACTGCTTGGAAATGCA 158
Db 692 AGTACTACCTTTCCATTTTATGGGATGAGTTCGTTGGGAGTCCCTTGGAAATACCA 751
QY 159 CTGTGTGTTGGCTACCTCTCTGATGATGAAGAACTGGAGCAAGCAGCATGCTATCTT 218
Db 752 TTGTTGTTTACGGCTACATCTCTCTCTGAAGAACTGGAGCAAGCAGCATGCTATCTT 811
QY 219 TTAACCTTTCCATCTCTGACTTTGCTTCTGTCACCTTCCCATCTCCATCTGATAAAGAGTT 278
Db 812 TTAACCTCTCTGCTCTGACTTACCTTTTCTGTCACCTTCCCATCTCCATCTGATAAAGAGTT 871
QY 279 ATGCCAATGATAAGGGACCTATGGAGATGTTCTCTGATAAGCAACCGATATGCTTC 338
Db 872 ATGCCAATGAACTGGATATATGGAGACGCTCTCTGATAAGCAACCGATATGCTTC 931
QY 339 ACACCAACCTCTACACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398
Db 932 ATGCCAACCTCTATACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 991
QY 399 TCATGAAGTACCTTTCCGAGAACACTTTCTTACAAAGAAAGAAATTTGCCATTTAACT 458
Db 992 TAATTAAGTATCTTTCCGAGAACACTTTCTGCAAAAGAAAGAGTTTGTCTATTTAACT 1051
QY 459 CGTGGCTGTCTGGGCTTTAGTACCTTAGAGTTCTACCCATCTGCTACCTTTTCATCAAT 518
Db 1052 CTTGGCCATTTGGGTTTGTAGTAACTTTAGTAACTTTAGTAACTTTAGTAACTTTAGTAACT 1111

Db 632 GGATCATGGCATGGAATGCAACTTGCAGAAACCTGGCTGGCAGCAGAGGCTGCCCTGGAA 691
QY 99 AGTACTACCTCTCTGCATTTATGCAATCGAGTTTCATTTTGGACTGCTTGGAAATGCA 158
Db 692 AGTACTACCTTTCCATTTTATGGGATGAGTTCGTTGGGAGTCCCTTGGAAATACCA 751
QY 159 CTGTGTGTTGGCTACCTCTCTCTGATGATGAAGAACTGGAGCAAGCAGCATGCTATCTT 218
Db 752 TTGTTGTTTACGGCTACATCTCTCTCTGAAGAACTGGAGCAAGCAGCATGCTATCTT 811
QY 219 TTAACCTTTCCATCTCTGACTTTGCTTCTGTCACCTTCCCATCTCCATCTGATAAAGAGTT 278
Db 812 TTAACCTCTCTGCTCTGACTTACCTTTTCTGTCACCTTCCCATCTCCATCTGATAAAGAGTT 871
QY 279 ATGCCAATGATAAGGGACCTATGGAGATGTTCTCTGATAAGCAACCGATATGCTTC 338
Db 872 ATGCCAATGAACTGGATATATGGAGACGCTCTCTGATAAGCAACCGATATGCTTC 931
QY 339 ACACCAACCTCTACACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398
Db 932 ATGCCAACCTCTATACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 991
QY 399 TCATGAAGTACCTTTCCGAGAACACTTTCTTACAAAGAAAGAAATTTGCCATTTAACT 458
Db 992 TAATTAAGTATCTTTCCGAGAACACTTTCTGCAAAAGAAAGAGTTTGTCTATTTAACT 1051
QY 459 CGTGGCTGTCTGGGCTTTAGTACCTTAGAGTTCTACCCATCTGCTACCTTTTCATCAAT 518
Db 1052 CTTGGCCATTTGGGTTTGTAGTAACTTTAGTAACTTTAGTAACTTTAGTAACTTTAGTAACT 1111

Db 519 CTGTCCCAAAAGAGAGGAGGAGTAACTGTCATGCACTGACATGCAAGTTCTCGAAACCCCTGAAC 578
Db 1112 CTGTTAATGATGCAATGGCACCACCTGTAATGATTTTGGCAAGTCTTGGAGACCCCAACT 1171
QY 579 ACAATCTCATTTAGAGCTCTGCTGACTTTGTTGGGCTTCCCTAACTCTCTCTCTCTGTA 638
Db 1172 ACAACCTCAITTTAGAGATGTTGTTAAACACTGTTGGGGTTTCCCTATTTCCCTTTTGTGA 1231
QY 639 TGTGCTTCTTCTACTACAAGATGTTAGTCTTCTTAAAGAGAGGAGGAGGAGGAGGAGGAG 698
Db 1232 TGTGTTCTTTTATACAGATGTTCTCTCTTCTTAAAGAGAGGAGGAGGAGGAGGAGGAG 1291
QY 699 CTGCCCTGCCACTGGCAAAACCCCAAGCCTGTTGTTGGGCTTCCCTAACTCTCTCTCTCTA 758
Db 1292 CTGCTCTGCCCTTTGAAAGCCTCTCAACTTGGTTCATCGGAGTGGTAACTCTCTCTCTG 1351
QY 759 TACTCTTACACACCTTATCATATATGCGCAATTTGAGGATCGGCTCAGCCTCGATAGATT 818
Db 1352 TGTCTTTTACACCTATGACGCTATGCGGATGTTGAGGATCGCTTACAGCCTGGGAGTT 1411
QY 819 G---GCCACAAGGATGTACAGAGGCCCAATCAAACTATATACACACTGACAGGGCCTC 875
Db 1412 GGAAGCAGTATCAGTGGCACTCAGGCTGCTCATCACTCTCTTACATTTGTGACAGGGCTT 1471
QY 876 TGGCCTTTCTGAACAGTGGCAATCAATCCCATCTTCTACTTCTCTCTCTCTCTCTCTCTCT 935
Db 1472 TGGGCTTTCTGAACAGTGTATCAACCCCTGCTCTCTATTTTCTTTTGGGAGATCACTTCA 1531
QY 936 GAGAGATGCTGATTAAGTTTACAGCAATACTTCAAGTCCCTTACATCTCTTACAGGACAT 995
Db 1532 GGGACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1591
QY 996 GAGCTGCTGGATGAGGCTTTTCACTACGCCAAA--TGAGACACTTTGATAAAGAC 1048
Db 1592 GGGCTCATGAACTCCTACTTTTCAATTCAGAGAAAGAGTGGGGGCTTGTGAAACAG 1645

RESULT 2
US-08-749-707-1
; Sequence 1, Application US/08749707
; Patent No. 6063582
; GENERAL INFORMATION:
; APPLICANT: Conley, Pamela B.

Db 703 GATGCCCGGCGAGAGCCTGTGGCCCGGAGGCGGCGTGGCAAGCGCGCGGCGATGGCCGTG 762
QY 737 CTGCGGTTGTGATCTTCTTACTACTTTCACACCCCTATCATATCATGCGCAATTGTAGG 796
Db 763 GTGGTGGCTGCTGCTTGGCCATCAGCTTCTGCGCTTTTACATACACCAAGACAGCCTAC 822
QY 797 ATGCGCTCAGCGCTGGATAGTTGGCCACAGGATGTACACAGAGCCCATCAATCTATA 856
Db 823 CTGGCAGTGGGCTCGAGCGCGGCGCTCCCTGCACTGTATTGGAGGCTTTGCAGCGGCC 882
QY 857 TACACACTGACAGCGCTCTGCGCTTTCGACAGTGGCCATCAATCCCATCTTCTACTTC 916
Db 883 TACAAGGCGACGCGGCGGCTTGGCAGTGCCACAGCGTGTGAGCCCATCTCTTCTAC 942
QY 917 CTCA 920
Db 943 TTCA 946

RESULT 7
US-08-012-988A-1
; Sequence 1, Application US/08012988A
; Patent No. 5652133
; GENERAL INFORMATION:
; APPLICANT: Murphy, Phillip M.
; TITLE OF INVENTION: Cloning and Expression of Human
; TITLE OF INVENTION: Macrophage Inflammatory Protein-1 alpha (MIP-1
; TITLE OF INVENTION: alpha)/RANTES Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/012,988A
; FILING DATE: 19930128
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-118
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2156 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 259..275
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: Complement (868..884)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 63..1128
; US-08-012-988A-1

Query Match 5.1%; Score 78.4; DB 1; Length 2156;
Best Local Similarity 54.1%; Pred. No. 3.1e-12;
Matches 160; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 379:
SEQUENCE CHARACTERISTICS:
LENGTH: 1023 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 37..1020
US-08-513-974B-379

Query Match 5.2%; Score 79.6; DB 3; Length 1023;
Best Local Similarity 46.0%; Pred. No. 9.3e-13;
Matches 388; Conservative 0; Mismatches 444; Indels 12; Gaps 3;

QY 89 ATCTTGAATGAAGTACTACCTCTCTGATTTTATGCAATCGAGTTCAATTTTGGACTGCTT 148
Db 103 AACTTCAAGCAACTGTGCTGCCACCTGTGATTCGGCGGTGCTGGCGCTGGCGCTGCG 162
QY 149 GGAATGCTACTGTGTGTTGGCTACCTCTTCTGCAATGAAGAACTGGAACAGCAGCAAT 208
Db 163 CTGNACATCTGTGCTATTACCCAGATCTGCACGTCGCGCGGCGCTGACCGCGCAGCGCC 222
QY 209 GTCTATCTTTTAACTTTCATCTCTGACTTTCCTTCTGACACCTTCCCTCCATCTG 268
Db 223 GTGTACACCTAAACCTGTCTGGCTGACCTGCTATATGCTGCTGCTGCTGCTGCTGCTC 282
QY 269 ATAAGAGTTATGCAAA---TGATAAGGGGACCTATGGAGATGTTCTGTATGAAGCAAC 325
Db 283 ATCTACAACTATGCCAAGGTGATCACTGCGCCTTTGGGACCTTGGCCTGGCGCTGGTC 342
QY 326 CGATATGCTGTACACCAACCTCTACACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCT 385
Db 343 CGCTTCTCTTCTATGCAACCTGCAAGGCGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCT 402
QY 386 GACGATATCTGCTCATGAAGTACCCCTTT---CCGAGAACACTTTCTACAAAAGAGAA 442
Db 403 CAGCGCTACCTGGGCGATCTGCCACCCCTGCGCCCTGGCCGCAACAGTGGGGCGCGCG 462
QY 443 TTTGCCATTTTAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502
Db 463 GCT 522
QY 503 CTCACCTTTCATCAATTTGTCCTCCCAAGAGAGGCGGAGTAACTGCACTGCACTATGCAAGT 562
Db 523 GCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 582
QY 563 TCTGGAACCTTGAACCACTCTCATTTACAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
Db 583 CTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
QY 623 ATTCCTCTCTCTGATGCTGCTTCTTCTACTACAAGATGTTAGTCTCTTAAAGAGGAGG 682
Db 643 CTGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702
QY 683 AGCCAGCAGCAAGCAACTG-----CCCTGCACTGGGCAACCCCAACCGCTGCTGCTGCT 736

us-09-891-138a-1.rni

2; Gaps 3;

TITLE OF INVENTION: Coupled Receptor GPR2
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 STREET: 1100 NEW YORK AVE., NW, SUITE 600
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,948A
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/04079
 FILING DATE: 30-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: STEFFE, ERIC K.
 REGISTRATION NUMBER: 36,688
 REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1301 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 161..1192
 PS-08-467-948A-7

Query Match 4.8%; Score 74.2; DB 2; Length 1301;
Best Local Similarity 49.9%; Pred. No. 3.6e-11;
Matches 187; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY	125	ATCGAGTTCATATTTTGGACTGCTTGGAATGTCAC	TGTGGTGTTCGGGTACCTCTTCTGCG	184
Db	236	ATGGTGTTCGTGGTAAATATCCAATTTGTTGGCATATACATTTTCATCTACGCGTC		295
QY	185	ATGAAGAACATGGAAACAGCAGCAATGTCATCTTTTAACTTTCCATCTCTCACATTTGCT	244	
Db	296	CTCAAGATCGGAATGAACATACAACTTACATGATTAACCTTGGCAATGTACAGACTTGCTT	355	
QY	245	TTCTGTGCACCTTCCCATCTTGATAAGAGGTTATGCCAATGATAAGGGGACCTATGGA	304	
Db	356	TTTGTTTTACTTTACGCTTCAGAGTTTTTACTTCCAAACGGAAATGGCCATTTGGA	415	
QY	305	GATCTTCTGTATTAAGCAACCGATATGTCTTTCACACCAACCTCTACACAGCATCTTC	364	
Db	416	GATTTACTTTGTAGATTTCTGTGATGCTGTTTTATACCAACATGTACGGAAGCAATCTG	475	
QY	365	TTCTCCTACTTTCAATAGCATGGACCGGATATCTGTCATGAAGTACCCCTTTCGAGAACAC	424	
Db	476	TTCTTAACCTGTATTAGTGTAGATTCGATTTCTGCGAAATTTGTCTACCCATTTAAGTCAAG	535	
QY	425	TTTCTCAAAAGAGGAATTTGCCATTTTAATCTCGCTGGCTCTCTCTGGGCTTTAGTGAAC	484	
Db	536	ACTCTAGAAGCAAAAGAAATGCAAGATTTGTTGCACTGGCGTGTGGTTAACTGTGATC	595	

RESULT 10
US-08-467-948A-7
; Sequence 7, Application US/08467948A
; Patent No. 5998164
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; APPLICANT: CAO, LIANG

QY 485 TTAGAAGTCTACCC 499
 Db 596 GGAGGAAGTGACCC 610

RESULT 11
 ; Sequence 7, Application US/08467947A
 ; Patent No. 6090575
 ; GENERAL INFORMATION:
 ; APPLICANT: LI, YI
 ; APPLICANT: CAO, LIANG
 ; APPLICANT: NI, JIAN
 ; APPLICANT: GENTZ, REINER
 ; APPLICANT: BULT, CAROL J.
 ; APPLICANT: SUTTON III, GRANGER G.
 ; APPLICANT: ROSEN, CRAIG A.
 ; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
 ; TITLE OF INVENTION: Coupled Receptor GPR1
 ; NUMBER OF SEQUENCES: 30
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVE., NW, SUITE 600
 ; CITY: WASHINGTON
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/467,947A
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/04079
 ; FILING DATE: 30-MAR-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: STEFFE, ERIC K.
 ; REGISTRATION NUMBER: 36,688
 ; REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-371-2600
 ; TELEFAX: 202-371-2540
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1301 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: both
 ; MOLECULE TYPE: cdna
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 161..1192
 ; US-08-467-947A-7

Query Match 4.8%; Score 74.2; DB 3; Length 1301;
 Best Local Similarity 49.9%; Pred. No. 3.6e-11;
 Matches 187; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY 125 ATCGAGTTCAATTTTGGACTGCTGGGAATGCTACTGTGGTGTGGCTACCTCTTCTGTC 184
 Db 236 ATGGGTGTTGTGGCTGGTAAATCAATGTTGCCATATACATTTTCATCTGGCTC 295

QY 185 ATGAAGAAGTGGACAGCAGCAAGTCTATCTTTTAACTTTCCATCTCTGACTTGTCT 244
 Db 296 CTCGAAGTCCCAATGAAGTCAACTTACATGATTAACCTGGCAATGTGACACTTGTCT 355

QY 245 TTCTGTGACCCCTTCCATCTCTGATAAAGAGTTATGCCAATGATGAAGGGGACCTATGGA 304

Db 356 TTGTTTTTACTTTACCCCTTACGAGATTTTACTTCACAACACGGAAATGGCCATTTGGA 415
 QY 305 GATGTTCTCTGTATAGCAACACGATATGTGCTTCACACCAACCTCTACACCAAGATCCTC 364
 Db 416 GATTACTTTGTAAGATTTCTGTGATGCTCTTTTATACCAACATGTACGAAGCATTTCTG 475

QY 365 TTCTCTACATTTTATAGCATGGACCATATCTGCTCATGAAGTACCCCTTTCCGAGAACAC 424
 Db 476 TTCTTAACCTGTATTAGTGTAGATGATTTCTGGCAATTTGCTACCCATTTAAGTCAAG 535

QY 425 TTCTCACAAGGAAGGAATTTGCAATTTAAATCTGCTGCTGTCTGGCCCTTAGTGACC 484
 Db 536 ACTCTAAGAACCAAAAGAAATGCAAGATTTTGTGCACTGGCGTGTAACTGTGATC 595

QY 485 TTAGAAGTCTACCC 499
 Db 596 GGAGGAAGTGACCC 610

RESULT 12
 ; Sequence 3, Application US/08383750
 ; Patent No. 5744301
 ; GENERAL INFORMATION:
 ; APPLICANT: Birkenbach, Mark
 ; APPLICANT: Kieff, Elliot
 ; TITLE OF INVENTION: Epstein Barr Virus Induced Genes
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 ; STREET: 1100 New York Avenue, N.W.,
 ; STREET: Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005-3934
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/383,750
 ; FILING DATE: Herewith
 ; CLASSIFICATION: 436
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fox, Samuel, L.
 ; REGISTRATION NUMBER: 30,353
 ; REFERENCE/DOCKET NUMBER: 0627.3300001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1643 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 34..1116
 ; US-08-383-750-3

Query Match 4.8%; Score 74; DB 1; Length 1643;
 Best Local Similarity 47.4%; Pred. No. 4.6e-11;
 Matches 255; Conservative 0; Mismatches 280; Indels 3; Gaps 1;

QY 126 TCGAGTTCAATTTTGGACTGCTGGGAATGTCACTGTGGTGTGGCTACCTCTTCTGCA 185
 Db 152 TCGTCTTCATCATCTGGCTGCTGGGAAACITACTAGCTTGGTGGTTCATTTGTTCAAAACA 211

QY 186 TGAACAACITGAACAGCAGCAATGCTATCTTTTAACTTTCCATCTCTGACTTTGCTT 245

FEATURE:
NAME/KEY: CDS
LOCATION: 34...1116
US-08-352-678-3

Query Match 4.8%; Score 74; DB 3; Length 1643;
Best Local Similarity 47.4%; Pred. No. 4.6e-11;
Matches 255; Conservative 0; Mismatches 280; Indels 3; Gaps 1;

126 TCGAGTTTCATTTTGGGACTGCTTGGGAATGCTCACTGCTGGGTGCTGGCTACCTCTCTCTGCA 185
152 TCGCTTTCATCATTTGGGCTGCTGGGAACCTTACTAGCCTTGGTGGCTGCTGTTCAAAACA 211
186 TGAAGAAGCTGGAACAGCAGCAATGCTCTCTTTTAAACCTTTCOATCTGACTTTGCTT 245
212 GGAATAAATCACTTACCAACCTCTATTCAACAAATTTGGTGATTTCTGATATACTTT 271
246 TCCTGTGACCCCTTCCCATCTCTGATAAAGAGTTATGCCAATG---ATAAGGGGACCTATG 302
272 TTACCACGGCTTTGCCCTACAGCAATAGCTTACTATGCAATGGGCTTTGACTGGAGAATCG 331
303 GAGATGTTCTCTGATATAAGCAACGATATGCTTACACCAACCTCTACACCAAGCATCC 362
332 GAGATGCCCTTGTAGGATAACTGCGCTAGTGTGTTTACATCAACACATATGACGCTGTA 391
363 TCTTCTCCTCCTTTCATTTAGCATGACCGATATCTGCTCATGAAGTACCTTTCCGAGAAC 422
392 ACTTTATGACCTGCTGAGTATGACCGCTTCAATGCTGTGCTGCTGCTGCTGCTGCTGCT 451
423 ACTTTCTCAAAAAGGAATTTGGCAATTTTAAATCTCGCTGGCTGCTGCTGCTGCTGCTGCT 482
452 ACAAGATAAAGGATTTGAACATGCAAAAGGGCTGTCATATTTGCTGCTGCTGCTGCTGCT 511
483 CTTTGAAGTTCTTACCCATGCTCACTTTTCATCAATTTCTGCTGCTGCTGCTGCTGCTGCT 542
512 TTGCTCAGACACTCCCACTCTCTCAACCTATGCTCAAAAGCAGGAGCTGAAAGGATTA 571
543 ACTGATCGACTATGCAAGTTCTGAAACCTGGAACACAACTCTCAATTTACAGCCTCTGCC 602
572 CATGATGGAGTATCCAACTTTGAAGAAGTAAATCTCTCCCTGGGATTTGCTGCTGGG 631
603 TGACTTTGTTGGGCTTCTCTTAATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
632 CATGTTTTCATAGCATATGCTTCCACTTATAATCAATTTCTCATCTGCTGCTGCTGCTGCT 689

RESULT 14
PCT-US93-09636-3
Sequence 3, Application PC/TUS9309636
GENERAL INFORMATION:
APPLICANT: Birkenbach, Mark
APPLICANT: Kieff, Elliott
TITLE OF INVENTION: Epstein Barr Virus Induced Genes
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: PCT/US93/09636
FILING DATE: herewith
CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540

212 GGAATAAATCAACTCTACCAACCTCTATTTCAACAAATTTGGTGATTTCTGATATACTTT 271
246 TCTGTGACCTTCCCTCCATCTGTATAAGAGTTATGCCAATG---ATAAGGGGACCTATG 302
272 TTACCACGGCTTTGGCTACAGNATAGCTTACTATGCAATGGGCTTTGACTGGAGAATCG 331
303 GAGATGTTCTGCTGATAAGCAACCGATATGCTTACACCAACCTCTACACCAACCTCC 362
332 GAGATGCCCTTGTAGGATAACTGCGCTAGTGTGTTTACATCAACACATATGACGCTGTA 391
363 TCTTCTCCTCCTTTCATTTAGCATGACCGATATCTGCTCATGAAGTACCTTTCCGAGAAC 422
392 ACTTTATGACCTGCTGAGTATGACCGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 451
423 ACTTTCTCAAAAAGGAATTTGGCAATTTTAAATCTCGCTGGCTGCTGCTGCTGCTGCTGCT 482
452 ACAAGATAAAGGATTTGAACATGCAAAAGGGCTGTCATATTTGCTGCTGCTGCTGCTGCT 511
483 CTTTGAAGTTCTTACCCATGCTCACTTTTCATCAATTTCTGCTGCTGCTGCTGCTGCTGCT 542
512 TTGCTCAGACACTCCCACTCTCTCAACCTATGCTCAAAAGCAGGAGCTGAAAGGATTA 571
543 ACTGATCGACTATGCAAGTTCTGAAACCTGGAACACAACTCTCAATTTACAGCCTCTGCC 602
572 CATGATGGAGTATCCAACTTTGAAGAAGTAAATCTCTCCCTGGGATTTGCTGCTGGG 631
603 TGACTTTGTTGGGCTTCTCTTAATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
632 CATGTTTTCATAGCATATGCTTCCACTTATAATCAATTTCTCATCTGCTGCTGCTGCTGCT 689

RESULT 13
US-08-352-678-3
Sequence 3, Application US/08352678
Patent No. 6043351
GENERAL INFORMATION:
APPLICANT: Birkenbach, Mark
APPLICANT: Kieff, Elliott
TITLE OF INVENTION: EPSTEIN BARR VIRUS INDUCED GENES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/352,678
FILING DATE: 30-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/980,518
FILING DATE: 25-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gates, Edward R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: B0801/7044
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 34..1116
PCT-US93-09636-3

```

```

Query Match          4.8%; Score 74; DB 5; Length 1643;
Best Local Similarity 47.4%; Pred. No. 4.6e-11;
Matches 255; Conservative 0; Mismatches 280; Indels 3; Gaps 1;

```

```

QY 126 TCGAGTTCATTTTGGAGTCTGGGAATGTCACCTGCTGGGTTCGGGTACTCTTTCGCA 185
DB 126 TCGAGTTCATTTTGGAGTCTGGGAATGTCACCTGCTGGGTTCGGGTACTCTTTCGCA 185
QY 152 TCGTTCATCATTTGGGCTCTGGGAACCTTACTAGCCTTGGTCTGATTTGTTCAAAACA 211
DB 152 TCGTTCATCATTTGGGCTCTGGGAACCTTACTAGCCTTGGTCTGATTTGTTCAAAACA 211
QY 186 TGAAGAACTGGAACAGCAGCAATGTCTATCTTTTAACTTTTCCATCTCTGACTTTGCTT 245
DB 186 TGAAGAACTGGAACAGCAGCAATGTCTATCTTTTAACTTTTCCATCTCTGACTTTGCTT 245
QY 212 GGAAGAACTGGAACAGCAGCAATGTCTATCTTTTAACTTTTCCATCTCTGACTTTGCTT 271
DB 212 GGAAGAACTGGAACAGCAGCAATGTCTATCTTTTAACTTTTCCATCTCTGACTTTGCTT 271
QY 246 TCCTGTCCACCTTCCCATCTGATGAAGAGTTATGCCAATG---ATAAGGGACCTATG 302
DB 246 TCCTGTCCACCTTCCCATCTGATGAAGAGTTATGCCAATG---ATAAGGGACCTATG 302
QY 272 TTACCACGGCTTTGGCTTACAGCAATAGCCCTACTGATGCAATGGGCTTTGACTGGGAATCG 331
DB 272 TTACCACGGCTTTGGCTTACAGCAATAGCCCTACTGATGCAATGGGCTTTGACTGGGAATCG 331
QY 303 GAGATGTTCTCTGTATGAAGCAACCGATATGTGCTTCCACCAACCTCTACACCAGCATCC 362
DB 303 GAGATGTTCTCTGTATGAAGCAACCGATATGTGCTTCCACCAACCTCTACACCAGCATCC 362
QY 332 GAGATGCTTGTGTAGGATTAACCTGCGCTAGTGTATTAACCAACATATGAGGTGTA 391
DB 332 GAGATGCTTGTGTAGGATTAACCTGCGCTAGTGTATTAACCAACATATGAGGTGTA 391
QY 363 TCCTTCTCCTTCTCCTGATGATGAGCGATATCTGCTATGAAGTACCTTTTCCGAGAAC 422
DB 363 TCCTTCTCCTTCTCCTGATGATGAGCGATATCTGCTATGAAGTACCTTTTCCGAGAAC 422
QY 392 ACTTTATGACCTGCTGAGTATGACGGCTTCTGCTGGTGGCACCCTCTAGCGTACA 451
DB 392 ACTTTATGACCTGCTGAGTATGACGGCTTCTGCTGGTGGCACCCTCTAGCGTACA 451
QY 423 ACTTCTTACAAAAGAGGATTTGCCATTTTATCTCGCTGGCTGCTGGGCTTCTAGTGA 482
DB 423 ACTTCTTACAAAAGAGGATTTGCCATTTTATCTCGCTGGCTGCTGGGCTTCTAGTGA 482
QY 452 ACAAGATAAAGAGATTGAACATGCAAAAGCGGTGTCATATTTGCTGGATTTCTAGTAT 511
DB 452 ACAAGATAAAGAGATTGAACATGCAAAAGCGGTGTCATATTTGCTGGATTTCTAGTAT 511
QY 483 CCTTGAAGTTTCCACCTGCTCCTTCTTCAATTTCTGCTCCCAAGAGAGGCGAGTA 542
DB 483 CCTTGAAGTTTCCACCTGCTCCTTCTTCAATTTCTGCTCCCAAGAGAGGCGAGTA 542
QY 512 TTGCTCAGACACTCCCACTCTCATCAACCTATGTCAAAGCAGGAGGCTGAAGGATTA 571
DB 512 TTGCTCAGACACTCCCACTCTCATCAACCTATGTCAAAGCAGGAGGCTGAAGGATTA 571
QY 543 ACTGCACTGCACTGCAAGTCTGGAACCCCTGGAACAAATCTCATTTACAGCCCTCTGCC 602
DB 543 ACTGCACTGCACTGCAAGTCTGGAACCCCTGGAACAAATCTCATTTACAGCCCTCTGCC 602
QY 572 CATGCACTGCACTGCAAGTCTGGAACCCCTGGAACAAATCTCATTTACAGCCCTCTGCC 631
DB 572 CATGCACTGCACTGCAAGTCTGGAACCCCTGGAACAAATCTCATTTACAGCCCTCTGCC 631
QY 603 TGACTTTGTTGGCTTCCTAAATTCCTCTCTCTGTGATGCTTCTTCTACTACAAGAT 660
DB 603 TGACTTTGTTGGCTTCCTAAATTCCTCTCTCTGTGATGCTTCTTCTACTACAAGAT 660
QY 632 CATGTTTCATAGGATATGATCTCCACTTATATCATTTCTCATCTGCTATTTCTCAGAT 689
DB 632 CATGTTTCATAGGATATGATCTCCACTTATATCATTTCTCATCTGCTATTTCTCAGAT 689

```

RESULT 15

```

US-08-847-296B-2
; Sequence 2, Application US/08847296B
; Patent No. 6271347

```

```

; GENERAL INFORMATION:
; APPLICANT: DAUGHERTY, BRUCE L.
; APPLICANT: DEMARTINO, JULIE A.
; APPLICANT: SICILIANO, SALVATORE J.
; APPLICANT: SPRINGER, MARTIN J.
; TITLE OF INVENTION: EOSINOPHIL EOTAXIN RECEPTOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

```

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,296B
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,158
; FILING DATE: 26-APR-1996
; APPLICATION NUMBER: 60/017,113
; FILING DATE: 26-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Thies, J. Eric
; REGISTRATION NUMBER: 35,382
; REFERENCE/DOCKET NUMBER: 19634Y
; TELEPHONE: 908-594-3904
; TELEFAX: 908-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1065 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-847-296B-2

```

```

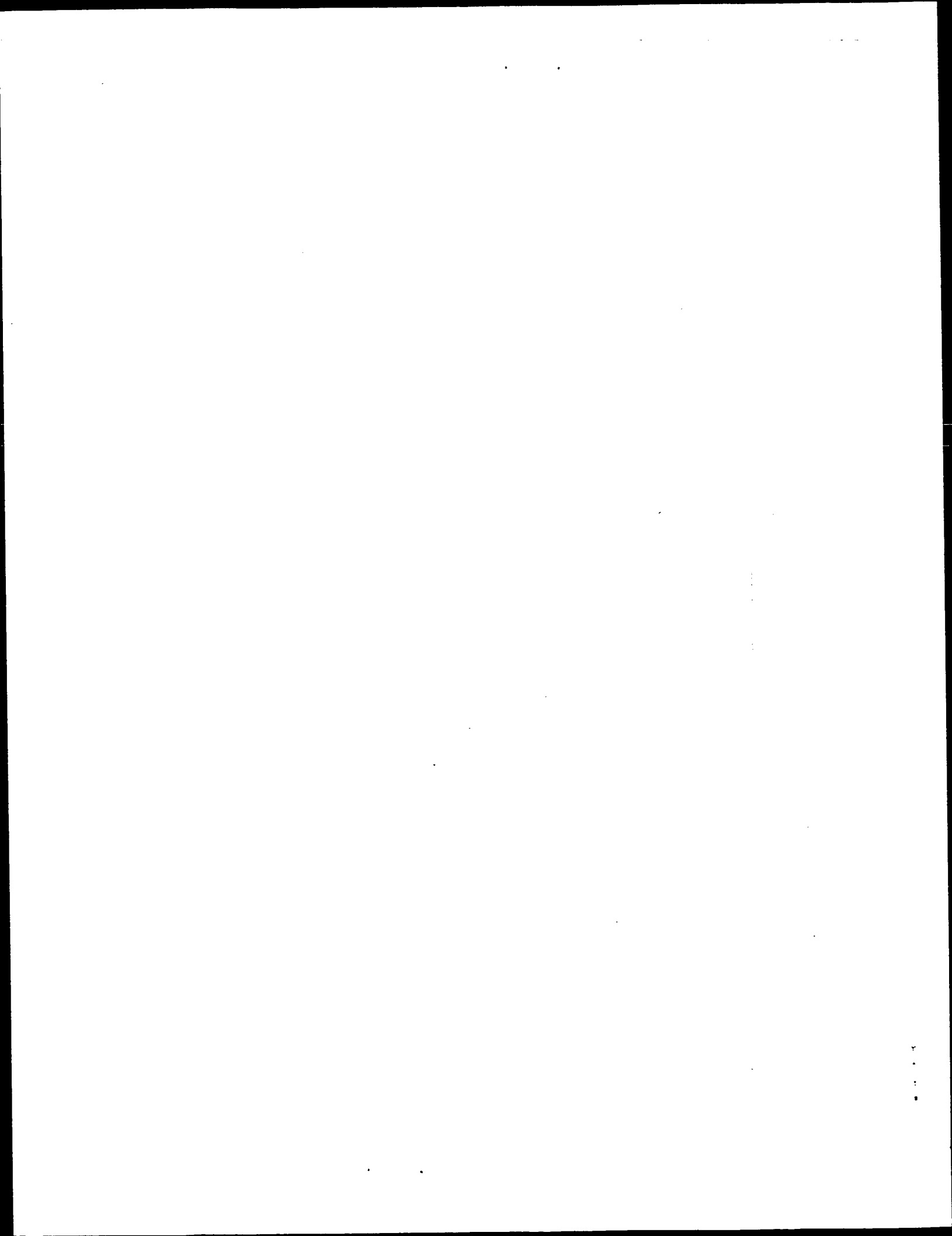
Query Match          4.8%; Score 73.8; DB 4; Length 1065;
Best Local Similarity 48.9%; Pred. No. 4.1e-11;
Matches 198; Conservative 0; Mismatches 207; Indels 0; Gaps 0;
QY 117 TTTATGCAATCGAGTTTCATTTTGGAGTCTTGGGAATGTCACCTGCTGGGTTCGGCTACC 176
DB 117 TTTATGCAATCGAGTTTCATTTTGGAGTCTTGGGAATGTCACCTGCTGGGTTCGGCTACC 176
QY 119 TGTATCTCCCTGCTGTTTCTGCTGCTGGGCTCTTGGGCAATGTTGGTGTGATGATCTCTCA 178
DB 119 TGTATCTCCCTGCTGTTTCTGCTGCTGGGCTCTTGGGCAATGTTGGTGTGATGATCTCTCA 178
QY 177 TCTTCTGCAATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTAACTTTTCCATCTCTG 236
DB 177 TCTTCTGCAATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTAACTTTTCCATCTCTG 236
QY 179 TAAATATACAGGAGGCTCCGAATTAAGCAACATCTACCTGCTCAACCTGGCCATTTGG 238
DB 179 TAAATATACAGGAGGCTCCGAATTAAGCAACATCTACCTGCTCAACCTGGCCATTTGG 238
QY 237 ACTTTGCTTTTCTGTCGACCTTCCCATCTGATGAAGAGTTATGCCAATGATGAAGGGA 296
DB 237 ACTTTGCTTTTCTGTCGACCTTCCCATCTGATGAAGAGTTATGCCAATGATGAAGGGA 296
QY 239 ACCTGCTCTTCTGCTGCTACCTTCCATTTGGATCCCACTATGTCAGGGGGCATAAATCTGG 298
DB 239 ACCTGCTCTTCTGCTGCTACCTTCCATTTGGATCCCACTATGTCAGGGGGCATAAATCTGG 298
QY 297 CCTATGAGAGTGTCTCTGATGAAGCAACCGATATGCTTTCACACCAACCTCTACACCA 356
DB 297 CCTATGAGAGTGTCTCTGATGAAGCAACCGATATGCTTTCACACCAACCTCTACACCA 356
QY 299 TTTTGGCCATGGCATGTGTAAGCTCTCTCAGGGTTTATATCACACAGGCTTGTACAGCG 358
DB 299 TTTTGGCCATGGCATGTGTAAGCTCTCTCAGGGTTTATATCACACAGGCTTGTACAGCG 358
QY 357 GCATCTCTTCTCTCCTTTCATTTAGCATGGACCGATATCTGCTCATGAAGTACCCCTTTC 416
DB 357 GCATCTCTTCTCTCCTTTCATTTAGCATGGACCGATATCTGCTCATGAAGTACCCCTTTC 416
QY 359 AGATCTTTTTCATAATCTGCTGACAAATCGACAGGTACCTGGCCATTTGCCATCTGTGT 418
DB 359 AGATCTTTTTCATAATCTGCTGACAAATCGACAGGTACCTGGCCATTTGCCATCTGTGT 418
QY 417 GAGAACACTTTTCTACAAAGAGGAATTTGCCATTTTAACTTCGCTGGCTGTCTGGGCT 476
DB 417 GAGAACACTTTTCTACAAAGAGGAATTTGCCATTTTAACTTCGCTGGCTGTCTGGGCT 476
QY 419 TTGCCCTTCGAGCCCGGACTGTCTACTTTTGGTGTGTCATCACACAGGCTTGTACAGG 478
DB 419 TTGCCCTTCGAGCCCGGACTGTCTACTTTTGGTGTGTCATCACACAGGCTTGTACAGG 478
QY 477 TAGTGACCTTTAGAAGTTTACCCCATGCTCACTTTTCATCAATTTCTG 521
DB 477 TAGTGACCTTTAGAAGTTTACCCCATGCTCACTTTTCATCAATTTCTG 521
QY 479 TGGCAGTCTAGCAGCTCTTCTCTGAATTTATCTTCTATGAGACTG 523
DB 479 TGGCAGTCTAGCAGCTCTTCTCTGAATTTATCTTCTATGAGACTG 523

```

```

Search completed: April 2, 2003, 15:13:46
Job time : 79 secs

```



Result No.	Score	Query Match	Length	ID	Description
1	590.8	38.3	1428	9 US-10-270-587-1	Sequence 1, Appli
2	126.6	8.2	1014	9 US-10-270-144-1	Sequence 1, Appli
3	126.6	8.2	1014	10 US-09-943-798-3	Sequence 3, Appli
4	126.6	8.2	9905	9 US-10-270-144-3	Sequence 3, Appli
5	125	8.1	1014	9 US-10-023-775B-1	Sequence 1, Appli
6	119	7.7	1313	10 US-09-728-422-1	Sequence 1, Appli
7	96	6.2	831	10 US-09-943-798-1	Sequence 1, Appli
8	82.2	5.3	1805	10 US-09-823-114-18	Sequence 18, Appli
9	82.2	5.3	1829	9 US-09-905-186A-9	Sequence 9, Appli
10	82.2	5.3	1829	9 US-09-905-186A-10	Sequence 10, Appli
11	82.2	5.3	1829	9 US-09-905-186A-11	Sequence 11, Appli
12	82.2	5.3	2534	9 US-10-087-345A-22	Sequence 22, Appli
13	80.6	5.2	1829	9 US-09-905-186A-7	Sequence 7, Appli
14	80.6	5.2	1829	9 US-09-905-186A-8	Sequence 8, Appli
15	80	5.2	1586	10 US-09-104-792-1	Sequence 1, Appli
16	77.6	5.0	1773	9 US-10-112-599A-3	Sequence 3, Appli
17	76.4	5.0	2602	9 US-09-905-186A-1	Sequence 1, Appli
18	74.2	4.8	1301	9 US-10-024-494-7	Sequence 7, Appli
19	73.8	4.8	1065	9 US-09-922-895-2	Sequence 2, Appli

Query Match	38.3%	Score 590.8	DB 9	Length 1428
Best Local Similarity	75.2%	Pred. No. 2.5e-135		
Matches 763	Conservative 0	Mismatches 247	Indels 4	Gaps 2
39 GCAGAATGGCAGACAATTTATCTGTGAGAATTTGGTTGGCAACAGAGGCTATCTTGAATA	98			
99				
99 GGATCATGGCATGGAATGCAACTTGCAAAACATGGCTGSCACAGAGGCTGCCCTGGAAA	158			
99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTGGACTGCTGGGAATGTCA	158			
159				
159 AGTACTACCTTTCATTTTATGGGATGAGTTCGTTGGGAGTCCTTGGAAATACCA	218			
159 CTGTGGTCTTCGGGTACCTCTCTGCATGAGACTGGCAACAGCAGCAATGTCTATCTTT	218			
219				
219 TTGTTGTTTACGGCTACATCTCTCTGAAAGACTGGAAACAGCAGTAAATTTATCTCT	278			
219 TTAACCTTTTCCATCTCTGACTTTCCTCTGCAACCTTCCCATCTGATAAAGAGTT	278			

Db	279	TTAAACCTCTCTGCTCTGACATTAGCTTTTGTGCAACCCCTCCCCATGCTGATGAAGGAGTT	338
Qy	279	ATGCCAATGATTAAGGGACCTTATGGAGATGTTCTCTGTATAAAGCAACCGGATATGTGCTTC	338
Db	339	ATGCCAATGGAACATGGATATATGGAGACGCTCTGCTATGAACCAACCGGATATGTGCTTC	398
Qy	339	ACACCAACCTCTACACGACATCCCTTCCTCACCTTTTCATTTAGCATGGACCGGATATCTGC	398
Db	399	ATGCCAACCTCTATACGACATCTCTTTTCACCTTTATCAGCATAGATCGATCTTGA	458
Qy	399	TCATGAAGTACCTTTCCGAGAACACTTTCTACAAAGAAGGAATTTGCCATTTTAATCT	458
Db	459	TAATTAGTATCCTTTCCGAGAACCTCTCGAAAGAAAGAGTGTGCTATTTTAATCT	518
Qy	459	CGCTGGCTGTCTGGGCTTTAGTGACCTTAGAAGTTCTACCCATGCTCACCTTTTCATCAATT	518
Db	519	CGTTGGCCATGTGGGTTTTAGTAACTTAGAGTTACTACCCATACTTCCCTTTATNAATC	578
Qy	519	CTGTCCCAAGAAGAGGGCAGTAAGTCATCGACTATGCAAGTTCTGGAACCCCTCAAC	578
Db	579	CTGTATACTGCAATGGCACCCTCTGAATGATTTTGCRAAGTTCTGGAGACCCCAACT	638
Qy	579	ACAATCTCATTTACGCTCTGCTGACCTTTCTGGGCTTCTCTATTCCTCTCTCTGTGA	638
Db	639	ACAACCTCATTTACAGCATGTCTTAACACTGTTGGGTTCTCTATTCTCTTTTGTGA	698
Qy	639	TGTGCTTCTTTTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGACGACGAGCAAGCA	698
Db	699	TGTGTTCTTTTATTACAAGATGGTCTCTCTCTTAAAGCAGAGGAATAGGCAGGTGCTA	758
Qy	699	CTGCCCTGCCACTGGACAAACCCACGCTGGTGGTCTGCGGGTTGTGATCTTCTCTA	758
Db	759	CTGCTCTGCCCTTGAAAGCCCTCTCAACTGGTCATCATGTCAGTGGTAACTTCTCTG	818
Qy	759	TACTCTTACACCCCTATCATATCATCGCAATTTGAGGATCGCCTCAGCCCTGGATGTT	818
Db	819	TGCTTTTTTACACCCCTATCATCGTCTCGGGAATGTGAGGATCGCTTCAGCCCTGGGAGTT	878
Qy	819	G - - -GCCACAGGATGTACACAGAAGGCCATAAATCTATATACACATGACACGGCCTC	875
Db	879	GGAAGCAGTATCATGTCGACTCAGGTGTCATCAACCTCTTTACATGTGTGACACGGCCTG	938
Qy	876	TGGCCTTTTGACAGTGCCATCAATCCCATCTTCTACTTCTCATGGGAGACCAATTACA	935
Db	939	TGGCCTTTCTGAACAGTGTATCAACCCCTGCTCTTATTTCTTGTGGGAGATCACTTCA	998
Qy	936	GAGAGATGCTGATTAGTAAAGTTCAGACAATPACTTCAAGTCCCTTACATCTTTCAGACAT	995
Db	999	GGGACATGCTGATGAATCACTGAGACACAACCTTCAATCCCTTACATCTTTAGCAGAT	1058
Qy	996	GAGCTGTGGATGAGGCTTCTACTCAGCCAAA - TGAGACACTTGATTAACAG	1048
Db	1059	GGGCTCATGAATCCTACTTTTCAGAGAAAAGTGAGGGGCTTGTGAACAG	1112

```

RESULT 2
US-10-270-144-1
; Sequence 1, Application US/10270144
; Publication No. US20030049790A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: RECEPTORS, AND USES THEREOF
; FILE REFERENCE: CL000750CON
; CURRENT APPLICATION NUMBER: US/10/270,144
; CURRENT FILING DATE: 2002-10-15
; PRIOR FILING DATE: 2000-05-18
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1014

```

	TYPE: DNA	ORGANISM: Human	US-10-270-144-1
	Query Match	8.2%; Score 126.6; DB 9; Length 1014;	
	Best Local Similarity	49.9%; Pred. No. 2.9e-21;	
	Matches 377; Conservative	0; Mismatches 369; Indels 9; Gaps	
QY	60	CTTGTGGAATTCGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTCTGCAATTTT	119
DB	59	CTTTTGGAAATTCGACTGATGAAGAACATCCCACCTCAAGATGCACCTACCTCCCTGTATTT	118
QY	120	ATGCAATTCGAGTTTCATTTTGGACTGCTTGGAAATCTCACTGTGGTGTTCGGCTACCTCT	179
DB	119	ATGGCATATCTTCTCGTGGGATTTCCAGGCAATCGAGTAGTATATCCACTTACATTT	178
QY	180	TCTGCATGAAGAACTGGAACAGACGAGAAATGCTATCTTTTACCTTTCCTCTCTGACT	239
DB	179	TCAAATGAGACCTTGGAGAGCAGCACCATCATATGCTGAACCTGGCGCTGCACAGATC	238
QY	240	TTGCTTTTCCGTGTGCACCTTCCCATCTCTGATAAGAGTTATGCAAT---GATAAGGGA	296
DB	239	TGCTGTATCTGACAGCCTCCCTTCTGATTCACCTACTATGCCAGTGGCGAAACTGGA	298
QY	297	CCTATGAGATGTTCTCTGATGAAGCAACCGATATGTGCTTACACCAACCTCTACACCA	356
DB	299	TCCTTGGAGATTCATGCTGAAGTTTATCCGCTTCAGCTTCCATTTCAACCTGTATAGCA	358
QY	357	GCATCCTCTCCTCACCTTCATTTAGCATGGACCGATATCTGCTCATGAAGTACCCCTTCC	416
DB	359	GCATCCTCTCCTCACCTGTTCAGCATCTTCGCGCTACTGTGTGATCATTCACCCAATGA	418
QY	417	GAGAACACTTCTACAAAAGAGAAATTCGCCATTTTAAATCTCGCTGGCTGTCTGGCCCT	476
DB	419	GCTGCTTTTCATTCACAAACTCGATGTCGAGTTGTAGCCTGTGCTGTGCTGGATCA	478
QY	477	TAGTGACCTTAGAAGTTCTACCCATGCTCAGTTTCATCATTAATCTGTCCCCAAGAGAGG	536
DB	479	TTTCACTGGTAGCTGCATTCGATGACCTTCTTGATCAGATCAACCAACAGGACCAACA	538
QY	537	CGAGTAAGTGCATCGACTATCGAAGTTCTGGAACCCCTGGAACACAATCTCATTTACAGCC	596
DB	539	GATCAGCCTGTCTGACCTCACCAAGTTCGG-----ATGAACCAATACTATTAAAGTGT	592
QY	597	TCTGCTGACTTTGTGGGCTTCCTAATCTCTCTCTGTGATGTGCTTCTCTACTACA	656
DB	593	ACAACCTGATTTTGACTCSAACTACTTTCTGCCCTCCCTTGGTGATAGTGACACTTTGCT	652
QY	657	AGATGGTAGTCTCTTAAGAGGAGGAGCCAGCAGCAAGCAACCTGCCTCGCCACTGGACA	716
DB	653	ATACACGATATTCACACTCTTGACCAATGACATGCAAACTGCACAGCTGCCTTAAGCAGA	712
QY	717	AACCCCAACGCTGTGTGCTTGGCGTTGTGATCTTCTCTATCTCTTACACCCCTATC	776
DB	713	AGACAGAGGCTAACCATCTTGCTACTCCTTGCAATTTTAGTATGTTTTTACCTTCC	772
QY	777	ATATCATGGCGAATTTGAGGATCGCCTCAGCCCTG	811
DB	773	ATATCTTGAGGGTCAATCGGATCGAATCTCGCTG	807

```

RESULT 3
; US-09-943-798-3
; Sequence 3, Application US/09943798
; Patent No. US20020065215A1
; GENERAL INFORMATION:
; APPLICANT: Glaxo Group Limited
; TITLE OF INVENTION: Polypeptide
; FILE REFERENCE: QG1021
; CURRENT APPLICATION NUMBER: US/09/943,798
; CURRENT FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0

```

; SEQ ID NO 3
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-943-798-3

Query Match
Best Local Similarity 8.2%; Score 126.6; DB 10; Length 1014;
Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;

```
QY 60 CTCTGTGAGATTTGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTCGCATTTT 119
Db 59 CTTTGGAAATGCACTGATGAAACATCCCACTCAAGATGCACCTACCTCCCTGTTATTT 118
QY 120 ATGCAATCGAGTTTCAATTTTGGGACTGCTTGGGAATGTCATCTGCTGGTTCGGTACTCT 179
Db 119 ATGCAATTTCTTCCCTGGGATTTCCAGGCAATGCACTAGTATGATATCCACTTACATTT 178
QY 180 TCTGATGAAGACTGGAACAGCAGCAATGTCTATCTTTTAAACCTTTCCATCTCTGACT 239
Db 179 TCAAAATGAGACCTTGGAGAGCAGCACCATCATTTATGCTGAACCTGGCTGCACAGATC 238
QY 240 TTGCTTTCCTGTGCACCTTCCCATCCCTGATRAAGATTATGCCAAT---GATAAGGGGA 296
Db 239 TGCTGTATCTGACACGCTCCCTTCCCTGATTCACCTTCCATTTCAACCTGTATAGCA 358
QY 297 CCTATGAGATGTTCTGTATGAAGCAACCGATATGCTTTCACACCAACCTCTACACCA 356
Db 299 TCTTTGGAGATTCATGTGTAAAGTTATCCGCTTCAGCTTCCATTTCAACCTGTATAGCA 358
QY 357 GCATCCTCTTCCACCTTTCATAGCATGAGCGGATATCTGCTCATGAAGTACCCCTTCC 416
Db 359 GCATCCTCTTCCACCTTTCATAGCATGAGCGGATATCTGCTCATGAAGTACCCCTTCC 416
QY 417 GAGAACACTTCTTACAAAAAGGAATTTGCCATTTTAACTTCGCTGGCTGCTGGCT 476
Db 419 GCTGCTTTTCCATTCACAACTCGATGTCAGTTGTCAGCTTGTAGCCTGTGCTGGTGTGATCA 478
QY 477 TAGTGACCTTAGAAGTTCTACCACTGCTCACTTTTCATCAATTTCTGTCCCAAGAGAGG 536
Db 479 TTTCACTGTAGCTGTTCATTTCCGATGACCTTCTTGATCATCAACCAAGAGGACCAACA 538
QY 537 GCAGTAACCTGCACTGACTATGCAAGTTCTGGAAACCTCGAACCAATCTCATTTACAGCC 596
Db 539 GATCAGCTGTCTGACCTCACCAGTTCCG-----ATGAACCTCAATACTATTAAGTGGT 592
QY 597 TCTGCTGACTTTGTTGGCTTCCTAATTCCTCTCTCTGATGTGCTTCTTCTACTACA 656
Db 593 ACAACCTGATTTTGACTGCAACTACTTTCTGCTCCCTTGTGTAGTACACTTTGCT 652
QY 657 AGATGCTAGTCTTTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 716
Db 653 ATACCAGGATTTTACACACTCTGACCCATGAGCTGCAAACTGACAGCTGCCTTAAGCAGA 712
QY 717 AACCCCAACGCTGTGCTGCTGGGGTGTGATCTCTCTATCTCTTCTACCTTCAACCCCTATC 776
Db 713 AAGCAGGAGGCTACCACTTCTGCTACTCTCTGCAATTTTACGTATGTTTTTACCCCTTCC 772
QY 777 ATATCATGCGCAATTTGAGGATCGCCTCACGCCCTG 811
Db 773 ATATCTTGAGGGTCATTCGGATCGAATCTCGCCTG 807
```

RESULT 4

US-10-270-144-3

; Sequence 3, Application US/10270144
; Publication No. US20030049790A1

GENERAL INFORMATION:

APPLICANT: WEI, Ming-Hui et al

TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED

TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR

TITLE OF INVENTION: PROTEINS, AND USES THEREOF

FILE REFERENCE: CL000750CON

; CURRENT APPLICATION NUMBER: US/10/270,144
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/205,196
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 9905
; TYPE: DNA
; ORGANISM: Human
US-10-270-144-3

Query Match

Best Local Similarity 8.2%; Score 126.6; DB 9; Length 9905;
Matches 377; Conservative 0; Mismatches 369; Indels 9; Gaps 2;

```
QY 60 CTTGTGAGATTTGGTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTCGCATTTT 119
Db 8367 CTTTGGAAATGCACTGATGAAACATCCCACTCAAGATGCACCTACCTCCCTGTTATTT 8426
QY 120 ATGCAATCGAGTTTCAATTTTGGGACTGCTTGGGAATGTCACCTGTGCTGGTTCGGCTACTCT 179
Db 8427 ATGGCAATTTCTTCCCTGGGATTTCCAGGCAATGCACTAGTATGATATCCACTTACATTT 8486
QY 180 TCTGATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTAAACCTTTCCATCTCTGACT 239
Db 8487 TCAAAATGAGACCTTGGAGAGCAGCACCATCATTTATGCTGAACCTGGCTGCACAGATC 8546
QY 240 TTGCTTTCCTGTGCAACCTTCCCATCTCTGATAAAGAGTTATGCCAAT---GATAAGGGGA 296
Db 8547 TGCTGTATCTGACCAAGCCTCCCTTCCCTGATTCACCTACTATGCCAGTGGCGAAACTTGA 8606
QY 297 CCTATGAGATGTTCTCTGTATGAAGCAACCGATATGCTTTCACCAACCTCTACACCA 356
Db 8607 TCTTTGGAGATTTCAATGTGTAAAGTTTATCCGCTTCAGCTTCCATTTCAACCTGTATAGCA 8666
QY 357 GCATCCTCTTCCCTCAGCTTTCATTTAGCATGAGCAGCATATCTCTCATGAAGTACCCCTTCC 416
Db 8667 GCATCCTCTTCCCTCAGCTTTCATTTAGCATGAGCAGCATATCTCTCATGAAGTACCCCTTCC 416
QY 417 GAGACACTTCTTCAAAAGAGGAATTTGCCATTTTAACTCTGCTGGCTGTCTGGGCT 476
Db 8727 GCTGCTTTTCCATTCACAAACTCGATGTGCAAGTTGTAGCTGTGCTGTGCTGGATCA 8786
QY 477 TAGTGACCTTAGAAGTTCTACCCATGCTCACTTTTCATCAATTTCTGTCCCAAGAGAGG 536
Db 8787 TTTCACTGCTAGCTGTCTATTCGCGATGACCTTCTTGATCATCAACCAAGAGGACCAACA 8846
QY 537 GCAGTAACCTGATCGACTATGCAAGTTCTTGGAAACCTGAAACCAATCTCATTTACAGCC 596
Db 8847 GATCAGCTCTCTGACCTCACCAGTTTCGG-----ATGAACCTCAATCTATTAAGTGGT 8900
QY 597 TCTGCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGATGTGCTTCTTCTACTACA 656
Db 8901 ACAACCTGATTTTGGTGGGCTTCCTAATTCCTCTCTCTGATGTGCTTCTTCTACTACA 656
QY 657 AGATGCTAGTCTTCTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 716
Db 8961 ATACCAGATTTATCCACACTCTGACCCATGAGCTGCAAACTGACAGCTGCCCTTAAGCAGA 9020
QY 717 AACCCCAACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 776
Db 9021 AAGCAGGAGGCTAACCACTTCTGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9080
QY 777 ATATCATGCGCAATTTGAGGATCGCCTCACGCCCTG 811
Db 9081 ATATCTTGAGGGTCATTCGGATCGAATCTCGCCTG 9115
```

RESULT 5

US-10-023-775B-1

; Sequence 1, Application US/10023775B

; Publication No. US2003002282A1

```

: GENERAL INFORMATION:
: APPLICANT: Pfizer Ltd. (EP(CB) only)
: APPLICANT: Pfizer Inc. (US, JP, EB except GB)
: APPLICANT: Fidock, Mark David
: TITLE OF INVENTION: No. US20030022282A1el Polypeptide
: FILE REFERENCE: PC10959AGPR
: CURRENT APPLICATION NUMBER: US/10/023,775B
: CURRENT FILING DATE: 2001-12-18
: PRIOR APPLICATION NUMBER: GB 0030854.4
: PRIOR FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: US 60/260,590
: PRIOR FILING DATE: 2001-01-09
: PRIOR APPLICATION NUMBER: US 60/296,660
: PRIOR FILING DATE: 2001-06-07
: PRIOR APPLICATION NUMBER: GB 0111031.1
: PRIOR FILING DATE: 2001-05-04
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1014
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-023-775B-1

```

Query Match	8.1%	Score 125;	DB 9;	Length 1014;
Best Local Similarity	49.8%;	Pred. No. 7.3e-21;		
Matches 376;	Conservative	0;	Mismatches 370;	Indels 9; Gaps
Qy	60	CTTGTGAGAAATGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACTCTCTCTGCAATTTT	119	
Db	59	CTTTTGGAAATTCGACTGATGAACATCCCACTCAAGATGCACCTACCTCGCTGTATTT	118	
Qy	120	ATGCAATPCGAGTTCAATTTTGGACTCCTTGGGAATGTCACTGTGTGTGTTTGGCTTACCTCT	179	
Db	119	ATGGCATTATCTTCCCTCGTGGGATTTCCAGGCAATGCAGTAGTGATATCCACTTACATTT	178	
Qy	180	TCTGTCATGAAGAACTGGAACAGACGACAAATGTCTATCTTTTAAACCTTTCCATCTCTGACT	239	
Db	179	TCAAAATGAGACCTTGGAGAGCAGCACCATCATTTATGCTGAACCTGGCGCTGCACAGATC	238	
Qy	240	TTGCTTTCTGTGCGACCTTCCCATCTCTGATGAAGATTATGCCAAT--GATAAGGGGA	296	
Db	239	TGCTGTATCTTGACAGCGTCCCTTCTGATTCACTACTACTTGCAGTGGCGAAATCGA	298	
Qy	297	CCTATGGAGATGTTCTCTGTATAAGCAACCGGATATGTGCTTCACACCAACCTCTACACCA	356	
Db	299	TCTTTGGAGATTCATGTGTAACTTTATCGGCTTCAGCTTCCATTTCAAACCTGTATAGCA	358	
Qy	357	GCATCTCTTCTCCTACCTTTTCATTAGCATGGACCGATATCTGCTCATGAAGTAGTACCCTTCC	416	
Db	359	GCATCCCTTCTCCTCACCTGTTTCAGCATCTCTCCGCTACTGTGTGATCATTCACCCAATGA	418	
Qy	417	GAGAACACTTCTACAAAAGGAAGAAATTTGCCATTTTAAATCTCGCTGGCTGTCTGGGCT	476	
Db	419	GCTGCTTTTCATTCAACAAACTCGATGTGCGAGTTGTAGCCGTGCTGTGGTGTGATCA	478	
Qy	477	TAGTGACCTTAGAAGTTCTTACCATGCTCACTTTTCATCAATCTGTGTCACAAAGAGAGG	536	
Db	479	TTTCACCTGGTAGCTGTCATTCCGATGACCTTCTTGATCATCATCAACACACAGGACCAACA	538	
Qy	537	CGAGTAAGTCACTGACATATCGAAGTTCTGGAACCTGAAACCAATCTCATTTACAGCC	596	
Db	539	GATCAGCCTGTCTCGACCTCACCAAGTTCCG-----ATGAACTCAATACTATTAGTGT	592	
Qy	597	TCTGCTGACTTTTGGGCTTCTTAATTCCTCTCTGTGTGATGTGCTTCTTCTACTACA	656	
Db	593	ACAACCTAATTTTGACTGCAACTACTTTCTGCTCCCTTTGGTGATAGTGACACTTTTGTCT	652	
Qy	657	AGATGGTAGTCTTCTTAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACA	716	
Db	653	ATACACAGATTATCCACACTCTGACCATCTGACATGCAAACTGACAGCTGCCTTAAGCAGA	712	
Qy	717	AAACCCAAACGCTCGTGTGCTGTGCGGTGTGTGATCTTCTATACCTTCTCAACCCCTATC	776	

```

Db      713  AAGCACGAAGGCTAAACATCTCTACTCCCTTGCATTTTACGTATGTTTTTACCCCTTC 777
QY      777  ATATCATCGCGCAATTTGAGGATCGCCTCACGCCTG 811
Db      773  ATATCTTGAGGGTCATTCGGATCGAATCTCGCCTG 807

RESULT 6
US-09-728-422-1
; Sequence 1, Application US/09728422
; Patent No. US20020128187A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No US20020128187A1el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 787C1P2F
; CURRENT APPLICATION NUMBER: US/09/728,422
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 1
; LENGTH: 1313
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (547)..(1239)
; US-09-728-422-1

```

Query Match	7.7%	Score 119;	DB 10;	Length 1313;
Best Local Similarity	52.3%;	Pred. No. 2.5e-19;		
Matches 288;	Conservative 0;	Mismatches 260;	Indels 3;	Gaps 1;
Qy 60	CTTTGGCAATTTGGTGGCAACAGAGGCTATCTTCAATAAGTACTACCTCTCTCGCATTTT	119		
Db 605	CTTTTGGAAATTCGACTGATGAAACATCCCACTCAAGATGCACCTCTCTGTTATTT	664		
Qy 120	ATGCAATCGATTTCATTTTGGACGCTTGGAAATGTCACCTGTGTGGCTACCTCT	179		
Db 665	ATGGCAATATCTTCCTCGGGATTTTCAGGCAATGCAGTAGTATCCACTTACATTT	724		
Qy 180	TCGCGATGAAGAACTGGAGACAGCAATGCTATCTTTTAACTTTCCATCTCTGACT	239		
Db 725	TCAAAATGAGACCTTGGAGAGCAGCACCATCATTATGCTGAACCTGGCTGCACAGATC	784		
Qy 240	TTGCTTTTCTGTGCAACCTTCCATCTCTGATTAAGAGTTATGCCAAT---	296		
Db 785	TGCTGTATCTGACCAAGCTCCCTCTCTGATTCACCTACTATCCAGTGGCGAAACTGGA	844		
Qy 297	CGATGGAGATGTTCTGTGTAAGCAACCGATATGTGCTTCACACCAACCTCTACACCA	356		
Db 845	TCCTTTGGAGATTCATGTGTAAGTTTATCCGCTTCAGCTTCCATTCACCTGTATAGCA	904		
Qy 357	GCATCCTCTCTCTCACTTTTCATTTAGCATGGACCGATATCTGCTCATGAAGTACCTTTCC	416		
Db 905	GCATCCTCTCTCTCACTTTTCAGCATGTTCCGCTACTTCCGCTACTGTGTGATCATTCACCCAAATGA	964		
Qy 417	GAGAACACTTCTTACAAAAAGAAATTTGCCATTTTAATCTCGCGTGTCTGGGCT	476		

Db 965 GCTGCTTTTCCATTACAAAACGATGTCAGTTGTAGCTGTGCTGTGGTGGATCA 1024
QY 477 TAGTGACCTTAGAGTTCTACCCAGCTCATTCTTCATCAATTCGTCCCAAAAGAGG 536
Db 1025 TTTCACTGGTAGCTGTCATTCGGATGACCTTCCTTGATCACATCAACCAAGGACCAACA 1084
QY 537 GCAGTAAGTGCATGACATGCAAGTTCTGTGAAACCCCTGAACACAATCTCATTTACAGCC 596
Db 1085 GATCAGCCTGTCTGACCTCACCAGTTCGGATGCACTAATCTAATTAAGTGGTACACC 1144
QY 597 TCTGCTGACT 607
Db 1145 TAATTTGACT 1155

RESULT 7
US-09-943-798-1
: Sequence 1, Application US/09943798
: Patent No. US20020065215A1
: GENERAL INFORMATION:
: APPLICANT: Glaxo Group Limited
: TITLE OF INVENTION: Polypeptide
: FILE REFERENCE: OG1021
: CURRENT APPLICATION NUMBER: US/09/943,798
: CURRENT FILING DATE: 2001-08-31
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 831
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-943-798-1

Query Match 6.2%; Score 96; DB 10; Length 831;
Best Local Similarity 49.4%; Pred. No. 8.8e-14;
Matches 311; Conservative 0; Mismatches 310; Indels 9; Gaps 2;

QY 185 ATGAAGACTGGGAACGACGACGATGCTATCTTTTAACTTTTCCATCTCTGACTTTGCT 244
Db 1 ATGAGACCTTGAAGACGACGACCATATGCTCAACTGGCTGACAGATCTGCTG 60
QY 245 TTCTCTGACACCTTCCCATCTGATAAGAGTATGCGCAAT---GATAGGGGACCTAT 301
Db 61 TATCTGACGACCTCCCTCTCTGATCTACTATGCGAGTGGCGAAACTGGATCTTT 120
QY 302 GGAGATGTTCTGTATAAGCAACCGATATGCTTTCACACCAACTCTACACGACGATC 361
Db 121 GGAGATTTTCATGTGAAGTTTATCCGCTTCAGCTTCCATTTCACTGTATAGCAGCATC 180
QY 362 CTCTTCTCATTTCATAGATGGACCGATATCTGCTCATGAAGTACCTTTCCCGAAG 421
Db 181 CTCTTCTCATTTCATAGATGGACCGATATCTGCTCATGAAGTACCTTTCCCGAAG 421
QY 422 CACTTTCTACAAAAGAGGAATTTGCAATTTTAACTCGCTGGCTGTGCGCCCTTAGTG 481
Db 241 TTTTCATTACAAAACACTGATGCGATGTCAGTTGTAGCTGTGCTGTGGTGGATCTTCA 300
QY 482 ACCTTAGAAGTTCTACCCATGCTCACTTTTCATCAATTTCTGCCAAAAGAGAGGCGAGT 541
Db 301 CTGGTAGCTGTCTATTCGATGACCTTCTTGATCATCAACCAACGACGACCAACAGATCA 360
QY 542 AACTGCATGACTATGCAAGTTCTGGAACCCCTGAACACAATCTCATTTACAGCCTTCGC 601
Db 361 GCCTGTCTGAGCCTCACCAGTTTCG-----ATGAACCTCAATACTATTAAGTGGTACAA 414
QY 602 CTGACTTTTGGGCTTCTTAATCTCTCTCTGTGTGATGCTCTTCTTCTACTACAAGATG 661
Db 415 CTGATTTTGACTGCACTACTTCTGCTCCCTTGGTGTAGTACACTTTGCTTATACC 474
QY 662 GTAGTCTTTTAAAGAGGAGGAGCCAGCAGCAAGCACTGCCCTGCCACTGGACAAACCC 721
Db 475 ACGATTATCCCACTCTGACCCATGAGCTGCAAACTGACACCTGCCCTTAAGCAGAAAGCA 534

QY 722 CAACGCGCTGTGCTCTGCGGGTGTGATCTCTCTATATCTCTACACCCCTATCATATC 781
Db 535 CGAAGCTAACCAATCTGCTACTCTCTGATTTTAGTATGTTTTTACCTTCCATATC 594
QY 782 ATCCGCAATTTGAGGATCGCTCACGCCG 811
Db 595 TTGAGGCTATTCGGATCGAATCTCGCTG 624

RESULT 8
US-09-823-114-18
: Sequence 18, Application US/09823114
: Patent No. US20020061554A1
: GENERAL INFORMATION:
: APPLICANT: EVANS, CHRISTOPHER J.
: KEITH, DUANE E.
: TITLE OF INVENTION: OPIOID RECEPTOR GENES
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORRISON & FOERSTER
: STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500
: CITY: WASHINGTON
: STATE: DC
: COUNTRY: USA
: ZIP: 20006-1888
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/823,114
: FILING DATE: 29-Mar-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/148,351
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: MURASHIGE, KATE H.
: REGISTRATION NUMBER: 29,959
: REFERENCE/DOCKET NUMBER: 22000-20526.22
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 887-1500
: TELEFAX: (202) 887-0763
: TELEX: 90-4030 MRSNFOERSWSH
: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1805 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 10..1119
: SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-823-114-18

Query Match 5.3%; Score 82.2; DB 10; Length 1805;
Best Local Similarity 44.5%; Pred. No. 3.4e-10;
Matches 379; Conservative 0; Mismatches 463; Indels 9; Gaps 1;

QY 85 GCCTATCTTGAATAAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCAATTTTGGACT 144
Db 147 GCCCTCGGCTCAAGTCAACCTCGTGGGGCTCTACTCGCGCTGTGTCTCGAGGGCT 206
QY 145 GCTTGGGAATGTCACTGTGTGTGCGGTACCTCTTCTGTCATGAAGAACTGGAACAGCAG 204
Db 207 CCTGGGAACCTGCTGTGTGTGTACGTACCTCTCAGGCACACCAAAATGAAGACGCCAC 266
QY 205 CAATGTCTATCTTTTAACTTTCCATCTCTGACTTTGCTTTCTCTGACCTTCCCAT 264
Db 267 CAATATTTTACATCTTTAACTTGGCCCTGGCCGACACTCTGGTCTCTGCTGACGCTGCCCTT 326

us-09-891-138a-1.rnpb

[illegible]

```
RESULT 13
US-09-905-186A-7
; Sequence 7, Application US/09905186A
; Publication NO. US20030008283A1
; GENERAL INFORMATION:
; APPLICANT: Kreek, Mary Jeanne
; APPLICANT: LaForge, Karl Steven
; TITLE OF INVENTION: Alleles for the Human Orphanin
; TITLE OF INVENTION: FQ/No. US20030008289A1
; TITLE OF INVENTION: Receptor Based Therapies for Treatment of
; Title of Invention: Methods of Treatment Based Thereon
```

```

; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/218,205
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1829
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-905-186A-7

Query Match          5.2%; Score 80.6; DB 9; Length 1829;
Best Local Similarity 44.4%; Pred. No. 8.4e-10;
Matches 378; Conservative 0; Mismatches 464; Indels 9; Gaps

-QY      85  GGCTATCTTGAATAAGACTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTGGACT 14
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db        171  GCCCTCTGGGCTCAAGGTCACCACTGTGGGGCTCTACCTGGCCCTGTGTCTCGAGGGGT 233

QY      145  GCTTGGGAATGTCACGTGTGGTGTTCGGCTACTCTCTCTGCAATGAAGAACTGGAACAGCAG 20
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db        231  CCTGGGGAACGTGCTTGTCTATGTAGCTATCTCTAGGCACACCAAAATGAAGACAGCCAC 29

QY      205  CAATGTCTATCTTTTAAACCTTTCCATCTCTGCATTTTCTTCTCTGTGCACCTTCCCAT 26
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db        291  CAATATTATCATCTTTAACTGGCCCTGGCCGACACTCTGGTCTCTGCAGCTGGCCCTT 35

QY      265  CCGTGAATAGAGTTATGCCAAATGATGAAGGGACCTTACGAGATCTCTCTGTATAAGCAA 32
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db        351  CCAGGGACAGCATCTCTCTGGGGCTTGTGGCGGTTTTGGGAATCGGTGTGCAAGACGT 41

```


ADDRESS: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
 ADDRESSEE: STUART & OLSTEIN
 STREET: 6 Becker Farm Road
 CITY: Roseland
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/104,792
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/461,244
 FILING DATE: 05-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Ferraro, Gregory D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/DOCKET NUMBER: 325800-445
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1586 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 431..1495
 US-09-104-792-1

Query Match 5.2%; Score 80; DB 10; Length 1586;
 Best Local Similarity 47.3%; Pred. No. 1.1e-09;
 Matches 276; Conservative 0; Mismatches 305; Indels 3; Gaps 1;

Qy	98	AAGTACTACCTCTCTGCAATTTTATGCAATCGAGTTCATTTTGGACTGCTGGGAATGTC	157
Db	533	AAGTTGCTCCTGCTGCTTTTATTGCTCTCTGCTTTGTTATTCAGTCTCTCGGAAACAGC	592
Qy	158	ACTGTGGTTCGGCTACCTCTCTGCAATGGAAGCTGGAACAGCAGCAATGCTATCTT	217
Db	593	CTGGTCATCCCTGGTCTCTGCTGCTGCAAGAGCTGAGGAGCATCAGATGTATACCTC	652
Qy	218	TTTAACCTTTCCATCTCTGACTTTGCTTTCTGTGACCCCTGCCATCCCTGATAAAGAT	277
Db	653	TTGAACCTGGCCCTGCTGACCTGCTTTTGTCTCTCTCTCCCTTCCAGACCTA---C	709
Qy	278	TATGCCAATGATAGGGACCTATGGAGATGTTCTCTGTATAGCAACCGATATGTGCTT	337
Db	710	TATCTGTGGACCATGGTGGTGTGGGACTGTAATGTGCAAGAGTGGTCTGGCTTTTAT	769
Qy	338	CACCAACCTCTACACGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	397
Db	770	TACATGGCTTCTACAGCAGCATGTTTTCATCACCCCTCATGAGTGTGGACAGGTACCTG	829
Qy	398	CTCATGAAGTACCCCTTTCCGAGAACACTTCTTACAAAAGAGGAATTTGCCATTTTAAATC	457
Db	830	GCTGTGTCATGCGTGTATGCTTAAAGGTGAGGAGCATCAGGATGGGCACAAACGCTG	889
Qy	458	TCGCTGGCTGTGGCTTGTAGTACCTTAGAAGTTCTTACCATGCTCTCTCTCTCTCTCT	517
Db	890	TGCCTGGCAGTATGCTTAACCGCATTTATGGCTACCATCCCATTTGCTAGTCTTTACCA	949
Qy	518	TCTGTCCCAAAAGAGAGGCGAGTAACTGTCATGCGACTATGCAAGTTCCTGGAAACCCGAA	577
Db	950	GTGGCTCTGAAGTGGTGTCTACAGTGTATTATTTTACATTTTACAAATCAACAGACTTTGAAG	1009

Qy 578 CACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTG 637
 Db 1010 TGGAGATCTTCACCAACTTCAAAATGAACATTTTAGGCTTGTGATCCCATTCACCATC 1069
 Qy 638 ATGTGCTCTCTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAG 681
 Db 1070 TTTATGTTCTGCTACATTTAAATCTCGACCATCTGAAGAGGTG 1113

Search completed: April 2, 2003, 15:16:37
 Job time : 170 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: April 2, 2003, 12:17:26 ; Search time 2695 Seconds
(without alignments)
9272.592 Million cell updates/sec

Title: US-09-891-138A-1
Perfect score: 1543
Sequence: 1 gctcctggcagagtttctg.....tgccataaataatcaatata 1543

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em.estba:*
- 2: em.esthum:*
- 3: em.estin:*
- 4: em.estmu:*
- 5: em.estov:*
- 6: em.estpl:*
- 7: em.estro:*
- 8: em.htc:*
- 9: gb.estl:*
- 10: gb.est2:*
- 11: gb.htc:*
- 12: gb.est3:*
- 13: gb.est4:*
- 14: gb.est5:*
- 15: em.estfun:*
- 16: em.estom:*
- 17: gb.gss:*
- 18: em.gss.hum:*
- 19: em.gss.inv:*
- 20: em.gss.pln:*
- 21: em.gss_vrt:*
- 22: em.gss.fun:*
- 23: em.gss.mam:*
- 24: em.gss.mus:*
- 25: em.gss.other:*
- 26: em.gss.pro:*
- 27: em.gss.rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	560	36.3	683	10	BB323771
2	495.8	32.1	520	9	AI663305
3	455	29.5	469	10	BB744515
4	438	28.4	458	10	BB746222
5	414	26.8	428	10	BB738743
6	403.8	26.2	422	10	BB847918

7	388.4	25.2	420	10	BB864882
8	384.8	24.9	426	10	BB778587
9	380.4	24.7	396	10	BB739482
C	363.8	23.6	367	9	AI649254
11	357.6	23.2	636	10	BB645274
12	354.2	23.0	416	10	BB846608
13	309.8	20.1	327	10	BB220946
14	296	19.2	877	12	BB402029
15	294.6	19.1	333	10	BB254869
16	283.2	18.4	323	10	BB220888
17	279	18.1	316	10	BB225749
18	275.2	17.8	326	10	BB500452
19	261.4	16.9	285	10	BB327439
C	255.8	16.6	633	17	AG083174
21	253.4	16.4	314	10	BB498575
22	241	15.6	310	10	BB221521
23	215.4	14.0	282	10	BB215653
24	214	13.9	312	10	BB498898
25	202	13.1	1101	17	CNS04VV3
26	199.8	12.9	226	10	BB213317
C	192.8	12.5	956	17	CNS028Y4
28	181	11.7	294	10	AW112068
C	158.8	10.3	639	10	AW612141
30	151	9.8	1026	17	CNS051MY
C	144.2	9.3	589	12	BF196066
32	141	9.1	1022	17	CNS04W90
C	135.8	9.1	582	10	BE221739
34	137.4	8.9	139	9	AI021184
35	90	5.8	638	9	AL675845
36	89.4	5.8	641	14	BQ396255
37	88.8	5.8	770	9	AJ452673
38	84.6	5.5	2542	11	AK017378
39	84.6	5.5	3001	11	AK005013
40	80.4	5.2	546	12	BG552112
41	78.4	5.1	952	14	BM917763
42	78.4	5.1	1074	14	BM917063
43	78.4	5.1	1167	14	BQ053936
44	77.8	5.0	623	14	BQ038875
45	76.8	5.0	672	10	BB638766

ALIGNMENTS

RESULT 1 BB323771 683 bp mRNA linear EST 31-AUG-2001
BB323771 RIKEN full-length clone B430012021 3', mRNA sequence.
LOCUS BB323771 RIKEN full-length clone B430012021 3', mRNA sequence.
DEFINITION Mus musculus cDNA clone B430012021 3', mRNA sequence.
ACCESSION BB323771 GI:15411432
VERSION BB323771.2
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 683)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,K., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT On Jul 11, 2000 this sequence version replaced gi:9032085.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222

Thu Apr 3 10:42:32 2003

Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>,
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapped-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
waghi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Matsui, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Yamanaka, I., Kiyosawa, H., Kondo, S., Saigo, T., Shinagawa, A., Aizawa
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T.,
Ishii, Y. and Hayashizaki, Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. *J. Struct.
Funct. Genomics* 3 pre, L72-L86 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

FEATURES
SOURCE

Location/Qualifiers

```

1. .683
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="B430012021"
/clone_lib="RIKEN full-length enriched, 4 days neonate
male adipose"
/sex="male"
/tissue_type="adipose"
/dev_stage="4 days neonate"
/lab_host="DH10B"
/notes="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory,
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA
primed with a primer [5',
GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length
cap-trapper. cDNA went through one round of normalized
to Rot = 10.0 and subtraction to Rot = 229.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGTTAAATAATTAATCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from La
p1C

```

BASE COUNT	226 a	125 c	117 g	215 t
------------	-------	-------	-------	-------

Query Match	36.3%	Score 560;	DB 10;	Length 683;
Best Local Similarity	98.7%;	Pred. No. 8.2e-125;		
Matches 596.	Conservative	0;	Mismatches 5;	Indels 3;
				Gaps 3;

943 CCTGATTAGTTACAGAC-AATACTTCAAG-TCCCTTACATCCTTC-AGGACATGAGC 999

D7 GCGGAGGGATTTTCTCAGTGACCCAAATAACTTCAGTTCCCTTACATGCCTTCAAGGACAATAGT 125
D8 CCGCATTTACTAAGTTCAGCCCAGGCCAAATACTTCAGTTCCCTTACATGCCTTCAAGGACAATAGT 125

DG 00 GCIGATTGGTAAAGTTACCCCTTCATCAGTAAGCAATCGGTGTTT
QY 1000 TGCTGGATGCAGGCTTCACTCAGCCAAAATGAGACACTTGATAAACAGCTGCTGTCAGT 1059

[illegible]

FEATURES

prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

mouse tissues.

FEATURES	SOURCE
1. <u>Age</u>	1. <u>Age</u>
2. <u>Gender</u>	2. <u>Gender</u>
3. <u>Marital Status</u>	3. <u>Marital Status</u>
4. <u>Education</u>	4. <u>Education</u>
5. <u>Income</u>	5. <u>Income</u>
6. <u>Occupation</u>	6. <u>Occupation</u>
7. <u>Religion</u>	7. <u>Religion</u>
8. <u>Political Affiliation</u>	8. <u>Political Affiliation</u>
9. <u>Health Status</u>	9. <u>Health Status</u>
10. <u>Travel History</u>	10. <u>Travel History</u>
11. <u>Employment Status</u>	11. <u>Employment Status</u>
12. <u>Family Size</u>	12. <u>Family Size</u>
13. <u>Home Ownership</u>	13. <u>Home Ownership</u>
14. <u>Vehicle Ownership</u>	14. <u>Vehicle Ownership</u>
15. <u>Insurance Status</u>	15. <u>Insurance Status</u>
16. <u>Subscription Services</u>	16. <u>Subscription Services</u>
17. <u>Charitable Donations</u>	17. <u>Charitable Donations</u>
18. <u>Volunteer Work</u>	18. <u>Volunteer Work</u>
19. <u>Real Estate Transactions</u>	19. <u>Real Estate Transactions</u>
20. <u>Banking History</u>	20. <u>Banking History</u>
21. <u>Investment Portfolio</u>	21. <u>Investment Portfolio</u>
22. <u>Travel Preferences</u>	22. <u>Travel Preferences</u>
23. <u>Food Preferences</u>	23. <u>Food Preferences</u>
24. <u>Exercise Habits</u>	24. <u>Exercise Habits</u>
25. <u>Shopping Habits</u>	25. <u>Shopping Habits</u>
26. <u>Media Consumption</u>	26. <u>Media Consumption</u>
27. <u>Technology Usage</u>	27. <u>Technology Usage</u>
28. <u>Environmental Concerns</u>	28. <u>Environmental Concerns</u>
29. <u>Community Involvement</u>	29. <u>Community Involvement</u>
30. <u>Personal Interests</u>	30. <u>Personal Interests</u>

BASE COUNT	150 a	75 c	82 g	151 t
ORIGIN				

Query Match	28.4%	Score 438;	DB 10;	Length 458;
Best Local Similarity	99.8%	Pred. NO. 2.6e-95;		
Conservative	0.	Mismatches	0;	Indels 1;
Gap	0.	Gaps	0;	Gaps 1;

Qy	1058	GTTCAGTCTTTAACTAAGTAAGCAACACCATCTTTAGGCTTTAGCTTTCACCATCCTCCAAAC	1117
Db	10	GTTCAGTCTTTAACTAAGTAAGCAACACCATCTTTAGGCTTTAGCTTTCACCATCCTCCAAAC	69
Qy	1118	CCCCAGGCTGGAGTACAAGCTCGGCTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGAT	1177
Db	70	CCCCAGGCTGGAGTACAAGCTCGGCTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGAT	129
Qy	1178	TTTAGGTTATACCCAGCAGTATCGGAAAAAATTAAGGCATGAGAAAAACATTGACATCTTCACT	1237
Db	130	TTTAGGTTATACCCAGCAGTATCGGAAAAAATTAAGGCATGAGAAAAACATTGACATCTTCACT	189
Qy	1238	TAGAAGCTGAACAAAAGAGAACAAATATTGTCAATGTTTGACACTTTAGGATCTGAATC	1297
Db	190	TAG - ACTGAACAAAGAGAACAAATATTGTCAATGTTTGACACTTTAGGATCTGAATC	248
Qy	1298	TTGGAATTTTAAGACCTCTTTTCTATCAGTGTAAAAGGAATACAGATAGCTAGTTGC	1357
Db	249	TTGGAATTTTAAGACCTCTTTTCTATCAGTGTAAAAGGAATACAGATAGCTAGTTGC	308
Qy	1358	AAATGCTGAATGCATTTCATCTTGCTCAGGTCGATAAGCGCTGTTTCTGGAATAGTCTTA	1417
Db	309	AAATGCTGAATGCATTTCATCTTGCTCAGGTCGATAAGCGCTGTTTCTGGAATAGTCTTA	368
Qy	1418	TTTTTATCTTGTAATTAAAAATTTATCTGAAAAAATGAATATTAATTCATGTGACACAT	1477

```

|||||
Db 369 TTTTATTCTGTAATATAAAATTTATGCGAAAGTGAATATATCAATGACACAT 428
Qy 1478 TAGATTCTCTATTGCAAAATATATTTCTT 1507
Db 429 TAGATTTCTATTGAAATATATATTTCTT 458

RESULT 5
BB738743 428 bp mRNA linear EST 15-OCT-2001
LOCUS BB738743 RIKEN full-length enriched, 6 days neonate spleen Mus
DEFINITION musculus cDNA clone F430109C18 3', mRNA sequence.
ACCESSION BB738743
VERSION BB738743.1 GI:161411748
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 428)
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
,A., Takahashi,F., Takaku-Akashira,S., Tanaka,T., Tomaru,A., Toya,I.,
Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1. 428
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clones="F430109C18"
/clone_lib="RIKEN full-length enriched, 6 days neonate
spleen"
/tissue_type="spleen"
/dev_stage="6 days neonate"
/note="vector: pSPORT1; Site:1: SalI; Site:2: NotI; This
clone is among a rearranged set of 15,247 clones from 11

```

```

embryo cDNA libraries (including preimplantation stage
embryos from unfertilized egg to blastocyst, embryonic
part of E7.5 embryos, extraembryonic part of E7.5 embryos
, and E12.5 female mesonephros/gonad) and one newborn
ovary cDNA library. Average insert size 1.5 kb. All
source libraries are cloned unidirectionally with Oligo(dT
)-Not primers. References include: (1) Genome-wide
expression profiling of mid-gestation placenta and embryo
using a 15,000 mouse developmental cDNA microarray, 2000,
Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
Large-scale cDNA analysis reveals phased gene expression
patterns during preimplantation mouse development, 2000,
Development, 127: 1737-1749; (3) Genome-wide mapping of
unselected transcripts from extraembryonic tissue of
7.5-day mouse embryos reveals enrichment in the t-complex
and under-representation on the X chromosome, 1998, Hum
Mol Genet 7: 1967-1978."
BASE COUNT 153 a 59 c 72 g 144 t
ORIGIN
Query Match 26.8%; Score 414; DB 10; Length 428;
Best Local Similarity 100.0%; Pred. No. 1.7e-89;
Matches 414; Conservative 0; Mismatches 0; Gaps 0;
Qy 1130 AGTACAAGCTGGTCCACATGATCAGAGGCGAGCTCTCTGTTCTGATTTTAGGTTATAC 1189
Db 1 AGTACAAGCTGGTCCACATGATCAGAGGCGAGCTCTCTGTTCTGATTTTAGGTTATAC 60
Qy 1190 CCAGAGTAGTGGAAAAAATAAGGCATGAGAAAGCATGACATCTTCACCTTAAGAAGCTGAAC 1249
Db 61 CCAGAGTAGTGGAAAAAATAAGGCATGAGAAAGCATGACATCTTCACCTTAAGAAGCTGAAC 120
Qy 1250 AAAGAGAGAACAAATATTCCTCAATGTTGGACACCTAGGATCTGAAATCTTGGAAATTTTA 1309
Db 121 AAAGAGAGAACAAATATTCCTCAATGTTGGACACCTAGGATCTGAAATCTTGGAAATTTTA 180
Qy 1310 AGACCTCTTTTCTATCATCTAGTCTAAAAGGAATACAGATAGTACTGCAATGCTGAATG 1369
Db 181 AGACCTCTTTTCTATCATCTAGTCTAAAAGGAATACAGATAGTACTGCAATGCTGAATG 240
Qy 1370 CATTTTCATCTGGTTCAGGTCGATGAGCGTGTCTGAAATAGTCTTATTTTATTTCTTG 1429
Db 241 CATTTTCATCTGGTTCAGGTCGATGAGCGTGTCTGAAATAGTCTTATTTTATTTCTTG 300
Qy 1430 TAATATTAATAATTTATGTGAAAAATGAATATAATTCATGTACACATTTAGATTTCTAT 1489
Db 301 TAATATTAATAATTTATGTGAAAAATGAATATAATTCATGTACACATTTAGATTTCTAT 360
Qy 1490 TTGAAATATPATTTTCTTGAAAAATAACTGCTGCTGCTAAATAAATCAATATA 1543
Db 361 TTGAAATATPATTTTCTTGAAAAATAACTGCTGCTGCTAAATAAATCAATATA 414

RESULT 6
BB847918 422 bp mRNA linear EST 26-NOV-2001
LOCUS BB847918 RIKEN full-length enriched, adult male kidney Mus musculus
DEFINITION cDNA clone F530201F11 5', mRNA sequence.
ACCESSION BB847918
VERSION BB847918.1 GI:17086293
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 422)
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
,A., Takahashi,F., Takaku-Akashira,S., Tanaka,T., Tomaru,A., Toya,T.

```



```

QY 1523 GTGCCCTAAATAATCAATATA 1543
      |||||||
Db 301 GTGCCCTAAATAATCAATATA 321

RESULT 14
BG402029 877 bp mRNA linear EST 12-MAR-2001
LOCUS 602466748F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4594810 5',
DEFINITION mRNA sequence.
ACCESSION BG402029 GI:13295477
VERSION 1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 877)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapps-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1336 row: p column: 11
High quality sequence stop: 542.
FEATURES
    source
        Location/Qualifiers
            1..877
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4594810"
                /clone_lib="NIH_MGC_75"
                /lab_host="DH10B (T1 phage-resistant)"
                /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
                SfiI (ggccattagggc); Site_2: SfiI (ggccattagggc); 5' and
                3' adaptors were used in cloning as follows: 5' adaptor
                sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence:
                5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
                C, or G and N = A, C, G, or T). Average insert size 1.65
                kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
                by PCR. This library was enriched for full-length clones
                and was constructed by Clontech Laboratories (Palo Alto,
                CA). Note: this is a NIH_MGC Library."
                /mol_weight=198 a
                /size=221 c
                /date="199 g
                /time=259 t

BASE COUNT 198 a 221 c 199 g 259 t
ORIGIN
Query Match 19.2%; Score 296; DB 12; Length 877;
Best Local Similarity 74.6%; Pred. No. 6.4e-61;
Matches 412; Conservative 0; Mismatches 135; Indels 5; Gaps 3;

QY 39 GCAGAAATGGCAGCAATTTATCTGTGAGAACTGGTGGCAACAGAGGCTATCTTGAATA 98
      |||||||
Db 108 GGATATGGCATGAATGCAACTTGCAGAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 167
      |||||||
QY 99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTTCATTTTGGACTCCTTGGGAATGCA 158
      |||||||
Db 168 AGTACTACCTTTCCATTTTATGGGATGTAGTTCGTTGGGAGTCTTGGAAATACCA 227
      |||||||
QY 159 CTGGGGTGTGGGTACCTCTCTGTCATGAGAACTGGCAACAGCAATGCTCTATCTTT 218
      |||||||
Db 228 TTGTTGTTTACGGGTACATCTCTCTCTGAAGAACTGGCAACAGCAATATTTATCTCT 287
      |||||||
QY 219 TTAACCTTTTCCATCTCTGACTTTTGTCTTCCTGTGACCCCTTCCCATCTCTGATAAAGAGTT 278
      |||||||
Db 288 TTAACCTCTCTCTGACTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 347
      |||||||

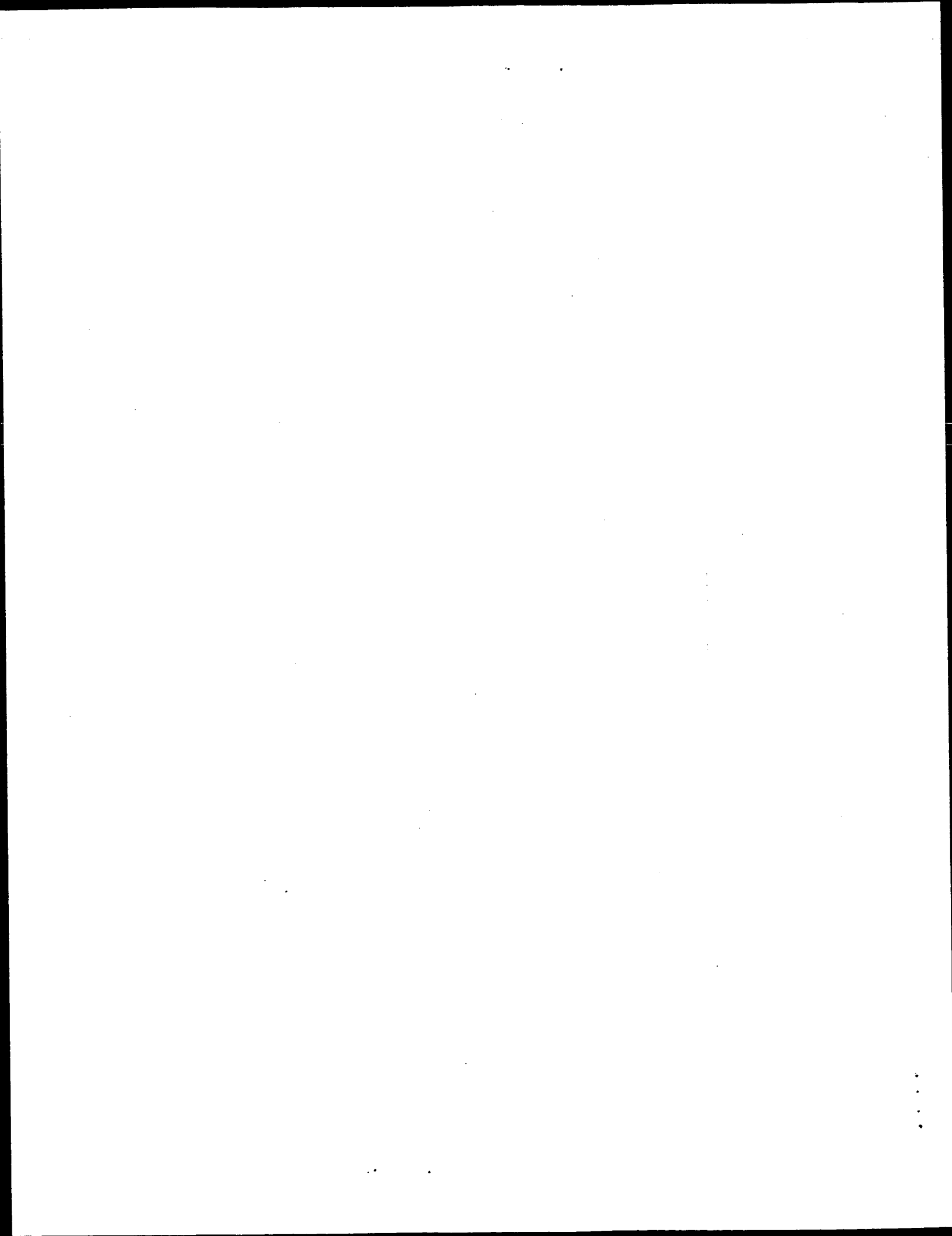
```

279 ATGCCAATGATAAGGGACCTATGGAGATGTTCTCTGTATATAAGCAACCGATATGCTCTC 338
|||||
348 ATGCCAATGGAACCTGGATATATGGAGACGCTGCTCTGCAATAAGCAACCGATATGCTCTC 407
|||||
339 ACAGCAACCTCTATACAGACATGCTCTCTCTCATTATGATAGCAGGACGATATCTGC 398
|||||
408 ATGCCAACCTCTATACAGACATGCTCTCTCTCATTATGATAGCAGGACGATATCTGC 467
|||||
399 TCATGAAGTACCTCTCCGAGAGACATTTCTACAAAAGAGGAATTTGCCATTTTAATCT 458
|||||
468 TAAATAAGTATCTCTCCGAGAGACACCTCTGCAAGAAAG--AGTGTCTATTTAATCT 525
|||||
459 CGCTGGCTGTCTGGGCCCTTAGTGACCTTAGAAGTTTACCCATGCTCCTCATCAAT 518
|||||
526 CCTGCC--ATTTGGTCTTAGTAACCTTAGAGTTTACCCATCTTCCCTTATAATC 583
|||||
519 CTGTCCCAAGAAGAGGGCAGTAACCTGCTGACTATGCAAGTTCTGGAACCCCTGAAC 578
|||||
584 CTGTTATTAACTGACAATGGCG--CCCTGTAAGGATTTTGAAGGCTGCGAGACCCCACT 642
|||||
579 ACAATCTCATTT 590
|||||
643 CCCACCTCTTTT 654

RESULT 15
BB254869
LOCUS BB254869
DEFINITION BB254869 RIKEN full-length enriched, 7 days neonate cerebellum Mus
musculus cDNA clone A730064B21 3', mRNA sequence.
ACCESSION BB254869
VERSION BB254869.1 GI:8947615
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 333)
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci-
P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
Y., Shigenoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya
T., Tsunoda,Y., Watahiki,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
Yano,R., Yasunishi,M., and Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sucho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoaka,S., Sasaki
N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermotabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95(2): 520-524 (1998)
Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res 9 (5): 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

1. *Chlorophyll a* (Chl *a*)

Search completed: April 2, 2003, 15:12:28
Job time : 2703 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 13:08:42 ; Search time 39 Seconds
(without alignments)
1083.088 Million cell updates/sec

Title: US-09-891-138a-2

Perfect score: 1650

Sequence: 1 MAQNSCENWLATEAILNKY.....REMLISKFRQVFKSLTSFRT 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1650	100.0	317	23	AAU74904
2	1238.5	75.1	334	21	AAU71308
3	1238.5	75.1	334	21	AAH02842
4	1238.5	75.1	334	23	ABB90381
5	1238.5	75.1	379	23	AAE15633
6	1236.5	74.3	334	18	AAW19854
7	1219.5	73.9	334	18	AAW22732
8	1216.5	73.7	387	22	AAU31029
9	1014.5	61.5	258	21	AAH45376
10	474	28.7	373	22	AAE04389

11	474	28.7	373	23	AAU10983	Purinergic recepto
12	474	28.7	373	23	AAU10984	Purinergic recepto
13	467	28.3	373	23	AAU10985	Purinergic recepto
14	426	25.8	337	22	AAU04375	Human G-protein co
15	426	25.8	337	23	AAU15399	Human G-protein co
16	426	25.8	337	23	ABH81902	Human G-protein co
17	426	25.8	337	23	ABH83819	Human G-protein co
18	426	25.8	337	23	ABH79438	Human P2Y-like rec
19	426	25.8	337	23	AAE21803	Human AXOR89 (G-pr
20	426	25.8	337	23	AAU77600	Human P2Y1-like Ho
21	426	25.8	337	23	AAU14027	Human P2Y1-like G
22	426	25.8	337	23	AAE16171	Human purinergic-r
23	422.5	25.6	374	22	AAE04390	Turkey P2Y nucleot
24	419	25.4	337	22	AAU04584	Human G-protein co
25	409	24.8	259	21	AAH45375	Gene 37 human secr
26	386.5	23.4	537	23	AAU74538	Human P2Y purinoce
27	379.5	23.0	336	22	AAH80971	Human NGPC854 #2
28	373.5	22.6	373	23	AAE20604	Mus musculus GPCR
29	366	22.2	276	23	ABH83818	Human P2Y-like rec
30	363.5	22.0	377	22	AAE04392	Human P2-purinergi
31	363.5	22.0	377	22	AAE01143	Human purinergic r
32	363.5	22.0	377	22	AAE01144	Human purinergic r
33	353	21.4	355	15	AAE53748	Seven transmembran
34	353	21.4	355	19	AAW48722	Human V28 seven tr
35	353	21.4	355	21	AAE21692	Human 7TM receptor
36	353	21.4	355	21	AAE21693	Human 7TM receptor
37	353	21.4	355	21	AAU90642	Human G protein-co
38	353	21.4	355	21	AAU90677	Human mutant G pro
39	353	21.4	355	22	AAH80126	Human CX3CR1 prote
40	353	21.4	355	22	AAH82786	Human CX3C chemoki
41	353	21.4	355	23	AAU91234	Human 7 transmembr
42	353	21.4	355	23	AAU91235	Human 7 transmembr
43	353	21.4	355	23	AAU84327	Protein CX3CR1 dif
44	352	21.3	365	18	AAW23606	Human P2Y4 recepto
45	352	21.3	365	22	AAE04391	Human P2Y4 pyrimid

ALIGNMENTS

RESULT 1

AAU74904
ID AAU74904 standard; Protein; 317 AA.

AC AAU74904;

DT 09-APR-2002 (first entry)

DE Amino acid sequence of mouse G-protein coupled receptor TGR18 protein.

KW Mouse; G-protein coupled; receptor; GPCR; TGR18; kidney disease;
KW signal transduction modulator; cerebral cavernous malformation;
KW hyperlipidemia; obesity; dyslexia; cardiac myxoma; renal failure;
KW nephritis; hypertension; liver disease; cirrhosis; blood disorder;
KW spleen-associated disorder; immune disorder.

OS Mus sp.

PN WO200200719-A2.

PP 03-JAN-2002.

PX 25-JUN-2001; 2001WO-US20363.

PR 23-JUN-2000; 2000US-213461P.

PA (TULA-) TULARIK INC.

PI Lin DC, Zhao J, Chen J, Cutler G;

DR WPI; 2002-147880/19.

XX N-PSDB; ABK12957.

PT	New G-protein coupled receptor polypeptides, useful for identifying	OS	Homo sapiens.
PT	modulators of signal transduction for treating kidney disease,	XX	WO200031258-A2.
PT	hyperlipidemia, obesity, dyslexia and cardiac myxoma	XX	02-JUN-2000.
XX	Claim 33; Page 59; 78pp; English.	XX	13-OCT-1999; 99WO-US23687.
XX	The present invention relates to a new G-protein coupled receptor (GPCR)	XX	20-NOV-1998; 98US-0109213.
CC	polypeptide comprising greater than 70% amino acid sequence identity to	XX	16-FEB-1999; 99US-0120416.
CC	the amino acid sequence of human GPCRs TGR62, TGR21, TGR130.1, TGR130.2,	XX	26-FEB-1999; 99US-0121852.
CC	human TGR213 or TGR92, 80% amino acid sequence identity to mouse TGR18	XX	12-MAR-1999; 99US-0123946.
CC	or 90% amino acid sequence identity to human novel edg receptor protein,	XX	12-MAR-1999; 99US-0123949.
CC	as defined in the specification. The GPCR covalently linked to a solid	XX	28-MAY-1999; 99US-0136436.
CC	phase is useful for identifying a compound that modulates signal	XX	28-MAY-1999; 99US-0136437.
CC	transduction. The identified compounds are useful for treating	XX	28-MAY-1999; 99US-0136439.
CC	kidney disease, cerebral cavernous malformations, hyperlipidemia,	XX	28-MAY-1999; 99US-0136567.
CC	obesity, dyslexia and cardiac myxoma. The molecules of the invention are	XX	28-MAY-1999; 99US-0137117.
CC	useful for diagnosing disorders or conditions such as kidney-related	XX	28-MAY-1999; 99US-0137131.
CC	conditions or diseases such as renal failure, nephritis, nephrotic	XX	29-JUN-1999; 99US-0141448.
CC	syndrome, asymptomatic urinary abnormalities, renal tubule defects,	XX	29-SEP-1999; 99US-0156555.
CC	e.g. cirrhosis, infiltrations, lesions, functional disorders and jaundice	XX	29-SEP-1999; 99US-0156633.
CC	and spleen-associated disorders or conditions e.g. splenic enlargement,	XX	29-SEP-1999; 99US-0156634.
CC	immune disorders, blood disorders and others. Modulation of the	XX	29-SEP-1999; 99US-0156653.
CC	polypeptide of the invention is useful to treat or prevent any of the	XX	01-OCT-1999; 99US-0157280.
CC	above conditions or diseases. The present amino acid sequence represents	XX	01-OCT-1999; 99US-0157281.
CC	the mouse GPCR TGR18 protein of the invention. This sequence is one of	XX	01-OCT-1999; 99US-0157282.
CC	seven novel G protein coupled receptors of the invention (AAU74904-	XX	01-OCT-1999; 99US-0157293.
CC	AAU74911).	XX	01-OCT-1999; 99US-0157294.
XX	Sequence 317 AA;	XX	12-OCT-1999; 99US-0416760.
XX	Query Match 100.0%; Score 1650; DB 23; Length 317;	XX	12-OCT-1999; 99US-0417044.
XX	Best Local Similarity 100.0%; Pred. No. 1.3e-157;	XX	(AREN-) ARENA PHARM INC.
XX	Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX	Chen R, Dang HT, Liaw CW, Lin I;
QY	1 MAQNLSNENLATEAILNKYLSAFYAIEFIFGLGNVTVVFGYLFCKMKNWSSNYLNF 60	XX	WPI: 2000-400068/34.
Db	1 MAQNLSNENLATEAILNKYLSAFYAIEFIFGLGNVTVVFGYLFCKMKNWSSNYLNF 60	XX	N-PSDB; AAD01135.
QY	61 LSISDFAFLCTPLILIKSYANDKGTGYDGLVCISNRYVLTNLTSLILFTFISMDRYLLM 120	PT	Novel human orphan G protein-coupled receptors and the encoding cDNAs
Db	61 LSISDFAFLCTPLILIKSYANDKGTGYDGLVCISNRYVLTNLTSLILFTFISMDRYLLM 120	PT	for use in the identification of G protein-coupled receptor agonists -
QY	121 KYPREHFLOKKEFALLISLAVWALVLEVLPLMTFTINSVPKEGSCIDYASSGNPEHN 180	XX	Claim 70; Page 87-88; 102pp; English.
Db	121 KYPREHFLOKKEFALLISLAVWALVLEVLPLMTFTINSVPKEGSCIDYASSGNPEHN 180	XX	The present amino acid sequence is the hCHN10, an endogenous human
QY	181 LYSISLCLTLLGLPLSLVMCFYFKMVVFLKRRSQQAATPLDLPORLVVLAIVFISIL 240	CC	orphan G protein-coupled receptor (GPCR), expressed in kidney and
Db	181 LYSISLCLTLLGLPLSLVMCFYFKMVVFLKRRSQQAATPLDLPORLVVLAIVFISIL 240	CC	thyroid. The hCHN10 cDNA was identified using the human EST (expressed
QY	241 FPPYHIMNRLRSLDSWPGCTQKATKSIYTLTRPLAFNSAINPIFYFLMGDHYREM 300	CC	sequence tag) 1365839 as a probe. The orphan GPCR of the invention, like
Db	241 FPPYHIMNRLRSLDSWPGCTQKATKSIYTLTRPLAFNSAINPIFYFLMGDHYREM 300	CC	all GPCRs has seven transmembrane alpha helices with an extracellular
QY	301 LISKFRQYFKSLTSPT 317	CC	N-terminus and an intracellular C-terminus. However, no endogenous
Db	301 LISKFRQYFKSLTSPT 317	CC	ligands has yet been identified for the proteins of the invention. The
XX	RESULT 2	CC	orphan GPCRs may be used in the identification of their endogenous
XX	AAU71308	CC	ligands, and to screen potential GPCR agonists and antagonists for use as
XX	AAU71308 standard; Protein: 334 AA.	CC	pharmaceutical agents. The proteins may also be used in the study of
XX	AAU71308;	CC	GPCR-mediated signalling cascades, and to elucidate their precise role in
XX	02-NOV-2000 (first entry)	CC	normal and diseased human conditions. Nucleic acid encoding human orphan
XX	Human orphan G protein-coupled receptor hCHN10.	CC	GPCRs may be used for tissue localisation expression analysis to provide
XX	Human; orphan G protein-coupled receptor; GPCR; hCHN10; drug screening;	CC	information about their function in healthy and pathological states.
KW	transmembrane receptor; expressed sequence tag; EST; signal cascade.	XX	Query Match 75.1%; Score 1238.5; DB 21; Length 334;
XX		XX	Best Local Similarity 72.2%; Pred. No. 3.6e-116;
XX		XX	Matches 228; Conservative 42; Mismatches 45; Indels 1; Gaps 1;
XX		QY	1 MAQNLSNENLATEAILNKYLSAFYAIEFIFGLGNVTVVFGYLFCKMKNWSSNYLNF 60
XX		Db	5 MAWNATCKNWLAAEALEKYLIFIFGIEFVGVGLNTIVVGYIFSLKNWSSNYLNF 64
XX		QY	61 LSISDFAFLCTPLILIKSYANDKGTGYDGLVCISNRYVLTNLTSLILFTFISMDRYLLM 120
XX		Db	65 LSISDFAFLCTPLILIKSYANDKGTGYDGLVCISNRYVLTNLTSLILFTFISMDRYLLM 124

QY 121 KYPREHFLQKKEFAILISLAWALVTLEVLPMLEINSVPKRGSCNIDYASSGNPEHN 180
 DB 125 KYPREHLLQKKEFAILISLAWALVTLEVLPMLEINSVPKRGSCNIDYASSGNPEHN 184
 QY 181 LIYSCLTLGLFLPLSVKMFYFYKMYVFLKRRSQOQATLPLDKPQRLVLAUVIFSIL 240
 DB 185 LIYSCLTLGLFLPLSVKMFYFYKMYVFLKRRSQOQATLPLDKPQRLVLAUVIFSIL 244
 QY 241 FTPYHMRNLRIASRLDSWPO-GCTQKAISYITLTPRLAFLNSAINPIFYFLMGDHYRE 299
 DB 245 FTPYHMRNLRIASRLDSWPO-GCTQKAISYITLTPRLAFLNSAINPIFYFLMGDHYRE 304
 QY 300 MLISKFRQYFKSLTSF 315
 DB 305 MLMNQLRHNFKSLTSF 320

RESULT 3
 AAB02842
 ID AAB02842 standard; Protein; 334 AA.
 AC AAB02842;
 XX
 DT 22-AUG-2000 (first entry)
 XX
 DE Human G protein coupled receptor hCHN10 protein SEQ ID NO:38.
 XX
 KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
 KW identification; agonist; screening; therapeutic; pharmaceutical;
 KW mutant.
 XX
 OS Homo sapiens.
 XX
 PN WO200022131-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 13-OCT-1999; 99WO-US24065.
 XX
 PR 13-OCT-1998; 98US-0170496.
 PR 12-NOV-1998; 98US-0108029.
 PR 20-NOV-1998; 98US-0109213.
 PR 27-NOV-1998; 98US-0110060.
 PR 16-FEB-1999; 98US-0120416.
 PR 26-FEB-1999; 98US-0121852.
 PR 12-MAR-1999; 98US-0123944.
 PR 12-MAR-1999; 98US-0123945.
 PR 12-MAR-1999; 98US-0123946.
 PR 12-MAR-1999; 98US-0123948.
 PR 12-MAR-1999; 98US-0123949.
 PR 12-MAR-1999; 98US-0123951.
 PR 28-MAY-1999; 98US-0136436.
 PR 28-MAY-1999; 98US-0136437.
 PR 28-MAY-1999; 98US-0136439.
 PR 28-MAY-1999; 98US-0137127.
 PR 28-MAY-1999; 98US-0137131.
 PR 30-JUN-1999; 98US-0137567.
 PR 27-AUG-1999; 98US-0141448.
 PR 03-SEP-1999; 98US-0151114.
 PR 29-SEP-1999; 98US-0152524.
 PR 29-SEP-1999; 98US-0156633.
 PR 29-SEP-1999; 98US-0156555.
 PR 29-SEP-1999; 98US-0156634.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
 XX
 DR WPI: 2000-317986/27.
 DR N-PSDB: AAA46036.
 XX
 PT Non-endogenous, human G protein-coupled receptors for screening

PT receptor, inverse or partial agonists useful as therapeutic agents -
 XX Example 1; Page 117-118; 187pp; English.
 XX The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCRs), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents,
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX
 XX Sequence 334 AA;

Query Match 75.1%; Score 1238.5; DB 21; Length 334;
 Best Local Similarity 72.2%; Pred. No. 3.6e-116;
 Matches 228; Conservative 42; Mismatches 45; Indels 1; Gaps 1;
 QY 1 MAONLSCENWLAETAILNKYYLSAFYAEIETIFGLGNVTVVFGYLFCKMKNSSNYLFN 60
 DB 5 MAWNAATCKNMLAAEALEKYYLSIFGYIEFVGVGLGNTIVVGYIFSLKNNSSNYLFN 64
 QY 61 LISIDFAFLCTLPILIKSYANDKGYDVLICISNRYLVHTNLTYSILFTFISMDRYLLM 120
 DB 65 LSVSDLAFLCTLPILIRSYANGNWIYGVLCISNRYLVHANLYTSILFTFISIDRYLII 124
 QY 121 KYPREHFLQKKEFAILISLAWALVTLEVLPMLEINSVPKRGSCNIDYASSGNPEHN 180
 DB 125 KYPREHLLQKKEFAILISLAWALVTLEVLPMLEINSVPKRGSCNIDYASSGNPEHN 184
 QY 181 LIYSCLTLGLFLPLSVKMFYFYKMYVFLKRRSQOQATLPLDKPQRLVLAUVIFSIL 240
 DB 185 LIYSCLTLGLFLPLSVKMFYFYKMYVFLKRRSQOQATLPLDKPQRLVLAUVIFSIL 244
 QY 241 FTPYHMRNLRIASRLDSWPO-GCTQKAISYITLTPRLAFLNSAINPIFYFLMGDHYRE 299
 DB 245 FTPYHMRNLRIASRLDSWPO-GCTQKAISYITLTPRLAFLNSAINPIFYFLMGDHYRE 304
 QY 300 MLISKFRQYFKSLTSF 315
 DB 305 MLMNQLRHNFKSLTSF 320

RESULT 4
 ABB90381
 ID ABB90381 standard; Protein; 334 AA.
 AC ABB90381;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 2757.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 XX neurological disease; infection; human; secreted protein.
 OS Homo sapiens.
 XX
 PN WO2001190304-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-US16450.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX

cell proliferative disorder; neurological; epilepsy; Parkinson's disease; Alzheimer's disease; inflammation; thyroiditis; haemolytic anaemia; AIDS; Acquired Immune Deficiency Syndrome; dementia; neurotic; cholelithiasis; multiple sclerosis; atherosclerosis; angina pectoris; gastroenteritis; diabetes; ulcer; viral infection; immunosuppressive.

Homo sapiens.

Location/Qualifiers

187..206

/label= Transmembrane_domain

234..253

/label= Transmembrane_domain

276..296

/label= Transmembrane_domain

319..342

/label= Transmembrane_domain

WO200198351-A2.

27-DEC-2001.

15-JUN-2001; 2001WO-US19275.

16-JUN-2000; 2000US-212483P.

22-JUN-2000; 2000US-213934P.

29-JUN-2000; 2000US-215209P.

07-JUL-2000; 2000US-216593P.

14-JUL-2000; 2000US-218936P.

19-JUL-2000; 2000US-219154P.

21-JUL-2000; 2000US-220141P.

(INCY-) INCYTE GENOMICS INC.

Lal P, Baughn MR, Hafalia AJA, Nguyen DB, Gandhi AR, Kallick DA; Griffin JA, Yue H, Khan FA, Patterson C, Lu DAM, Tribouley CM; Lu Y, Walia NK, Graul R, Yao MG, Yang J, Rankumar J, Au-Young J; Elliott VS, Hernandez R, Walsh RT, Borowsky ML, Thornton M, He A; WPI: 2002-075627/10. N-PSDB; AAD24958.

Isolated human G-protein coupled receptor polypeptides and the use of these sequences in the diagnosis, treatment and prevention of diseases and in the assessment of exogenous compounds on the expression of the receptors -

Claim 1; Page 115-116; 143pp; English.

The invention relates to isolated human G-protein coupled receptor (GCRC) polypeptides and their biologically active fragments. GCRC and protein is useful in treating a disease or condition associated with an increase or decrease in expression of functional GCRC. The GCRC's are useful in the diagnosis, treatment and prevention of cell proliferative disorders (cancer, leukaemia, melanoma); neurological disorders (stroke, epilepsy, Parkinson's disease, dementia, Alzheimer's disease); autoimmune inflammatory disorder (thyroiditis, haemolytic anaemia, AIDS, multiple sclerosis); cardiovascular disorder (atherosclerosis, angina pectoris); gastrointestinal disorder (ulcer, cholelithiasis, gastroenteritis); metabolic disorders (diabetes); viral infections (herpes virus) and in the assessment of the effects of exogenous compounds on the expression of the nucleic acid and amino acid sequences. The present sequence is human GCRC-3 protein.

Query Match 75.1%; Score 1238.5; DB 23; Length 379;

Best Local Similarity 72.2%; Pred. No. 4.3e-116;

Matches 228; Conservative 42; Mismatches 45; Indels 1; Gaps 1;

Qy 1 MAONLSCENLATEALNKYLLSAFYAIEFIFGLGNVTWVFYFCMKNNSSNYLNF 60

Db 50 MAWNAACKNLAAEALEKYLIFGYIEFVGVGLNTIVVYIFSLKNNSSNYLNF 109

Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders -

Claim 11; SEQ ID NO 2757; 2081pp + Sequence Listing; English.

The invention relates to novel genes (ABL89449-ABL90853) and proteins (AB899040-AB899044) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WPI at ftp.wipo.int/pub/published_pct_sequences.

Sequence 334 AA;

Query Match 75.1%; Score 1238.5; DB 23; Length 334;

Best Local Similarity 72.2%; Pred. No. 3.6e-116;

Matches 228; Conservative 42; Mismatches 45; Indels 1; Gaps 1;

Qy 1 MAONLSCENLATEALNKYLLSAFYAIEFIFGLGNVTWVFYFCMKNNSSNYLNF 60

Db 5 MAWNAACKNLAAEALEKYLIFGYIEFVGVGLNTIVVYIFSLKNNSSNYLNF 64

Qy 61 LSTDSFAFLCTLPILKSVANDKGTGVDVLCISNRYVLTNLTYSILFTLSMDRVL 120

Db 65 LSVSDLAFLCTLPILKSVANDKGTGVDVLCISNRYVLTNLTYSILFTLSMDRVL 124

Qy 121 KYPREHFLQKKEFALLISLVNVALVTLVPLMTIFINSVPKEGSGNCDYASSGNPEHN 180

Db 125 KYPREHFLQKKEFALLISLVNVALVTLVPLMTIFINSVPKEGSGNCDYASSGNPEHN 184

Qy 181 LIYSCLTLGLFLIPVSMCFYFYKMWVFLKRRSQOQATALPDKPQRLVLAIVFSIL 240

Db 185 LIYSCLTLGLFLIPVSMCFYFYKMWVFLKRRSQOQATALPDKPQRLVLAIVFSIL 244

Qy 241 FTHPHMRNLRIASRLDSWPQ-CCTOKAKTSIYTLRPLAFNLNSAINTFFYFLMGDHYRE 299

Db 245 FTHPHMRNLRIASRLDSWPQ-CCTOKAKTSIYTLRPLAFNLNSAINTFFYFLMGDHYRE 304

Qy 300 MLISKERQYFKSLTSF 315

Db 305 MLMQLRHNFKSLTSF 320

RESULT 5

AAE15633

ID AAE15633 standard; Protein; 379 AA.

XX AC AAE15633;

XX DT 12-MAR-2002 (first entry)

XX DE Human G-protein coupled receptor-3 (GCRC-3) protein.

XX KW Human; G-protein coupled receptor-3; GCRC-3; therapy; cancer; stroke;


```

QY 61 LSISDAFLCTPLILIKSYANDKGTGYDVLICISNRVYVLTNLYTSILFLTFISMDRYLLM 120
Db 110 LSVSDLAFLCTPLMLIRSYANGNWIYGDVLCISNRVYVLTNLYTSILFLTFISMDRYLLI 169
QY 121 KYPREHFLQKKEFAILISIAVWLVLEPLMLTFINSVPKKEGSCNDIYASSGNPEHN 180
Db 170 KYPREHLLQKKEFAILISIAVWLVLEPLMLTFINSVPKKEGSCNDIYASSGNPEHN 229
QY 181 LIYSCLTLGLFLPLSVCMCFYKVMVFLKRSQQOATALPLDKPQRLVWLVAVIFSIL 240
Db 230 LIYSWCLTLGLFLPLSVCMCFYKVMVFLKRSQQOATALPLDKPQRLVWLVAVIFSIL 289
QY 241 FTPYHMRNRIASRLDSWPQ-GCTQKAIKSIYTLTRPLAFNLNSAINPIFYELMGDHYRE 299
Db 290 FTPYHVMNRVRIASRLGSKWQYQCTQVINSFVIVTRPLAFNLNSAINPIFYELMGDHYRE 349
QY 300 MLISKFRQYKSLTSF 315
Db 350 MLMNQLRHNFKSLTSF 365

RESULT 6
AAW19854
ID AAW19854 standard; Protein; 334 AA.
XX AC AAW19854;
XX DT 11-SEP-1997 (first entry)
XX DE Human purinergic receptor P2U2.
XX KW P2U2 receptor; purinergic receptor; diagnosis; therapy.
XX OS Homo sapiens.
XX FH Key
XX FT Location/Qualifiers
XX FT 1..23
XX FT Domain
XX FT 24..49
XX FT Domain
XX FT /label= TMI
XX FT /note= "transmembrane domain I"
XX FT 50..60
XX FT Domain
XX FT /label= ICDI
XX FT /note= "intracellular domain I"
XX FT 61..82
XX FT Domain
XX FT /label= TMII
XX FT /note= "transmembrane domain II"
XX FT 83..99
XX FT Domain
XX FT /label= ECDI
XX FT /note= "extracellular domain I"
XX FT 100..119
XX FT Domain
XX FT /label= TMIII
XX FT /note= "transmembrane domain III"
XX FT 120..141
XX FT Domain
XX FT /label= ICDII
XX FT /note= "intracellular domain II"
XX FT 142..161
XX FT Domain
XX FT /label= TMDIV
XX FT /note= "transmembrane domain IV"
XX FT 162..183
XX FT Domain
XX FT /label= ECDII
XX FT /note= "extracellular domain II"
XX FT 184..207
XX FT Domain
XX FT /label= TMDV
XX FT /note= "transmembrane domain V"
XX FT 208..233
XX FT Domain
XX FT /label= ICDIII
XX FT /note= "intracellular domain III"
XX FT 234..256
XX FT Domain
XX FT /label= TMDVI
XX FT /note= "transmembrane domain VI"
XX FT 257..276

```

```

FT FT /label= ECDIII
FT FT /note= "extracellular domain III"
FT FT 277..300
FT FT /label= TMDVII
FT FT /note= "transmembrane domain VII"
FT FT 301..334
FT FT /label= C-terminal_domain
XX XX WO9720045-A2.
XX XX 05-JUN-1997.
XX XX 08-NOV-1996; 96WO-US18175.
XX XX 15-NOV-1995; 95US-0559524.
XX XX 15-NOV-1995; 95US-0006782.
XX XX (CORT-) COR THERAPEUTICS INC.
XX XX Conley PB, Jantzen H;
XX XX WPI; 1997-310601/28.
XX XX N-PSDB; AAT71900.
XX XX New isolated purinergic receptor sub-type - used to develop
XX XX products for diagnosis and therapy, e.g. for screening for agonists
XX XX and antagonists which can modulate activation
XX XX Claim 1; Fig 1A-B; 36pp; English.
XX XX P2U2 receptor (AAW19854) is a novel human purinergic receptor
XX XX subtype that is abundantly expressed in kidney and in many cell
XX XX lines of megakaryocytic or erythroleukemic origin and which is
XX XX activated by ATP, UDP, UTP and UDP. Its amino acid sequence was
XX XX deduced from a cDNA clone derived from DAMI (ATCC CRL 9792) cells.
XX XX P2U2 and its polypeptides can be expressed in host cells and used
XX XX to develop diagnostic and therapeutic agents. Antagonists and
XX XX agonists based on the extracellular domains of P2U2 receptor, or
XX XX which affect receptor function by binding to one of the
XX XX intracellular domains, can be used to treat diseases caused by
XX XX aberrant activation of this receptor or to treat diseases whose
XX XX symptoms can be ameliorated by stimulating or inhibiting the
XX XX activity of the receptor.
XX XX Sequence 334 AA;
XX XX
XX XX Query Match 74.3%; Score 1226.5; DB 18; Length 334;
XX XX Best Local Similarity 71.5%; Pred. No. 5.8e-115;
XX XX Matches 226; Conservative 42; Mismatches 47; Indels 1; Gaps 1;
QY 1 MAONLSCENWLATAILNKYVLSAFYATIEFIFGLGNVTYVFGYLCMKWNSSNYLFN 60
Db 5 MAWNATCKNWLAAEPALEKYYLSIFYGIEFVGVGLNTIYVGYIFSLKWNSSNYLFN 64
QY 61 LSISDAFLCTPLILIKSYANDKGTGYDVLICISNRVYVLTNLYTSILFLTFISMDRYLLM 120
Db 65 LSVSDLAFLCTPLMLIRSYANGNWIYGDVLCISNRVYVLTNLYTSILFLTFISMDRYLLI 124
QY 121 KYPREHFLQKKEFAILISIAVWLVLEPLMLTFINSVPKKEGSCNDIYASSGNPEHN 180
Db 125 KYPREHLLQKKEFAILISIAVWLVLEPLMLTFINSVPKKEGSCNDIYASSGNPEHN 184
QY 181 LIYSCLTLGLFLPLSVCMCFYKVMVFLKRSQQOATALPLDKPQRLVWLVAVIFSIL 240
Db 185 LIYSWCLTLGLFLPLSVCMCFYKVMVFLKRSQQOATALPLDKPQRLVWLVAVIFSIL 244
QY 241 FTPYHMRNRIASRLDSWPQ-GCTQKAIKSIYTLTRPLAFNLNSAINPIFYELMGDHYRE 299
Db 245 FTPYHVMNRVRIASRLGSKWQYQCTQVINSFVIVTRPLAFNLNSAINPIFYELMGDHYRE 304
QY 300 MLISKFRQYKSLTSF 315
Db 305 MLMNQLRHNFKSLTSF 320

```

RESULT 7	
AAW22732	
ID AAW22732 standard; Protein; 334 AA.	
XX	
AC AAW22732;	
DT 07-OCT-1997 (first entry)	
XX	
DE Human ATP receptor.	
XX	
KW ATP receptor; G-protein coupled receptor; agonist; antagonist.	
XX	
OS Homo sapiens.	
PH Key	Location/Qualifiers
FT Misc-difference 212	/note= "encoded by TCC"
FT Misc-difference 235	/note= "encoded by TCG"
FT Misc-difference 244	/label= Unknown
FT	/note= "encoded by CYT"
XX:	
PN WO9724929-A1.	
XX	
PD 17-JUL-1997.	
XX	
PF 11-JAN-1996; 96WO-US00392.	
XX	
PR 11-JAN-1996; 96WO-US00392.	
XX	
PA (HUMA-) HUMAN GENOME SCI INC.	
XX	
PI Li Y;	
XX	
DR WPI; 1997-372505/34.	
XX	
DR N-PSDB; AAT75146.	
XX	
PT Isolated human ATP receptor - agonists and antagonists of which are	
PT useful in treatment of, e.g. asthma, hypertension, arterial	
PT thrombosis and psychotic and neurological disorders	
XX	
PS Claim 15; Fig 1A-C; 53pp; English.	
XX	
CC Human ATP receptor (AAW22732) is structurally related to the G	
CC protein-coupled receptor family. It shows 29.8% identity to a	
CC murine P2U receptor. Its amino acid sequence was deduced from a	
CC human placental cDNA clone (AAT75146). Recombinant ATP receptor can	
CC be expressed in bacterial (e.g. E. coli), mammalian (e.g. COS) or	
CC insect (e.g. Sf9) host cells and used to screen for agonists and	
CC antagonists useful in the treatment of conditions related to	
CC underexpression of the receptor (e.g. asthma, Parkinson's disease,	
CC acute heart failure, hypotension, urinary retention and	
CC osteoporosis) or underexpression of the receptor (e.g. arterial	
CC thrombosis, hypertension, thrombolysis, angioplasty, cystic	
CC fibrosis, ulcers, asthma, allergy, benign prostatic hypertrophy,	
CC psychotic and neurological disorders, dyskinesias, endogenous	
CC anorexia and bulimia).	
XX	
SQ Sequence 334 AA;	
	Query Match
	Best Local Similarity 73.9%; Score 1219.5; DB 18; Length 334;
	Matches 225; Conservative 43; Mismatches 47; Indels 1; Gaps 1;
QY 1 MAQNLSNENLATEAILNKYILSAFYATEFTGLGNVTGVYGLFCMKNWSSNVYLEN 60	
DB 5 MAWNATCNWLAEEAALEKYLISFYTGIEFVVGVLGNTFIVVYGFISLKNWSSNIVLEF 64	
QY 61 LSISDFALCTLPILIKSYANDKGTGVDVLCISNRYVLTNLTYSILFTFISMDRYLLM 120	

SO	Sequence	387 AA;	
Query Match	73.7%;	Score 1216.5;	DB 22; Length 387;
Best Local Similarity	71.1%;	Pred. No. 7.2e-114;	
Matches	224;	Conservative 43;	Mismatches 47; Indels 1; Gaps 1;
QY	2 AONLSCNLIATRAILNKKYLSAFYAEIFGGLGNVTVVFGYLFCKMKNWSSNYVLENL	61	
DB	59 AWAATCKHLLAAEALEKYYLSIFGYIEFVGVLGNTIVVGYIFSLKNWSSNYVLENL	118	
QY	62 SISDFAPLCTPLILIKSYANDKTYGDVLCISNRYVHLNLTYSILFTFTSMRDYLLMK	121	
DB	119 SVSDLAFLCTPLMLIRSYANDNMIYGDVLCISNRYVHLNLTYSILFTFTSIDRYLLIK	178	
QY	122 YPREHFLOKKEFAILLISLAVLWLVLELPLMTFTNSVPKREGSNCIDYASSGNPEHL	181	
DB	179 YPREHLLQKKEFAILLISLWLVLELPLMTFTNSVPKREGSNCIDYASSGNPEHL	238	
QY	182 IYSCLTLLGLFLPLSVMGFFYKVVFLKRSQQOATLPLDKPQRLVLAIVFISILF	241	
DB	239 IYSCLTLLGLFSIPLFVCMFLFYKIALFKQNRQVATLPLEKPLNLVIMAVIFSILF	298	
QY	242 TPVHIMRNRIASRLDSWPQ-GCTQKAISYITLTPRPLAFINSAINPIFYFLMGDHYREM	300	
DB	299 TPVHYMRNRIASRLGSKWQYQCTQVINSFYIVTRPLAFINSVNPVYFELLDGDFRDM	358	
QY	301 LISKFRQYFKSLTSF 315		
DB	359 LMQLRHNFKSLTSF 373		
RESULT 9			
AA045376			
ID	AA045376 standard; Protein; 258 AA.		
AC	AA045376;		
XX	14-FEB-2001 (first entry)		
DT	Human secreted protein sequence encoded by gene 37 SEQ ID NO:128.		
DE	Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;		
XX	antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;		
KW	cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;		
KW	fungicide; ophthalmological; vulnerable; gene therapy; autoimmune disease;		
KW	hyperproliferative disorder; cardiovascular disorder; angiogenesis;		
KW	cerebrovascular disorder; nervous system disorder; infection; skin aging;		
OS	ocular disorder; wound healing; food additive; preservative.		
OS	Homo sapiens.		
PN	WO2000061628-A1.		
XX	19-OCT-2000.		
PD	06-APR-2000; 2000WO-US09070.		
PF	09-APR-1999; 99US-0128695.		
PR	14-JAN-2000; 2000US-0176052.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
PA	Rosen CA, Ruben SM, Komatsoulis G;		
XX	WPI; 2000-619228/59.		
DR	New nucleic acid molecules encoding 49 human secreted proteins for		
PT	diagnosing, preventing, treating or ameliorating medical conditions and		
PT	used as food additives or preservatives -		
XX	Disclosure; Page 447-448; 454pp; English.		
PS	The polynucleotide sequences given in AAC81086 to AAC81134 encode the		
XX			
CC			
CC	human secreted proteins given in AAB45308 to AAB45356. AAB45357 to		
CC	AAB45384 represent human secreted polypeptide sequences and proteins		
CC	homologous to them, which are given in the exemplification of the present		
CC	invention. Human secreted proteins have activities based on the tissues		
CC	and cells the genes are expressed in. Examples of activities include:		
CC	antiarthritic; immunosuppressive; antirheumatic; antiproliferative;		
CC	cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;		
CC	neuroprotective; antibacterial; virucide; fungicide; ophthalmological;		
CC	and vulnary. The polynucleotides and polypeptides can be used to		
CC	prevent, treat or ameliorate a medical condition in e.g. humans, mice,		
CC	rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used		
CC	in diagnosing a pathological condition or susceptibility to a		
CC	pathological condition. Disorders which are diagnosed or treated include		
CC	autoimmune diseases, hyperproliferative disorders, cardiovascular		
CC	disorders, cerebrovascular disorders, angiogenesis, nervous system		
CC	disorders, infections caused by bacteria, viruses and fungi and ocular		
CC	disorders. The polypeptides can also be used to aid wound healing and		
CC	epithelial cell proliferation, to prevent skin aging due to sunburn, to		
CC	maintain organs before transplantation, for supporting cell culture of		
CC	primary tissues, to regenerate tissues and in chemotaxis. The		
CC	polypeptides can also be used as a food additive or preservative to		
CC	increase or decrease storage capabilities, fat content, lipid, protein,		
CC	carbohydrate, vitamins, minerals, cofactors and other nutritional		
CC	components. AAC81077 to AAC81085 and AAB45307 represent sequences used in		
CC	the exemplification of the present invention.		
XX	Sequence 258 AA;		
SQ			
Query Match	61.5%;	Score 1014.5;	DB 21; Length 258;
Best Local Similarity	72.9%;	Pred. No. 9e-94;	
Matches	188;	Conservative 33;	Mismatches 36; Indels 1; Gaps 1;
QY	20 YVLSAFYAEIFGGLGNVTVVFGYLFCKMKNWSSNYVLENLSDFAFLCTPLILIKSY 79		
DB	1 YVLSFYGIEFVGVLGNTIVVGYIFSLKNWSSNYVLENLSDFAFLCTPLILIKSY 60		
QY	80 ANDKGTGYDVLCSNRYVHLNLTYSILFTFTSMRDYLLMKYPFHEHLOKKEFALLIS 139		
DB	61 ANGNLYGDVLCSNRYVHLNLTYSILFTFTSIDRYLLIKYPFHEHLOKKEFALLIS 120		
QY	140 LAVWLVLELPLMTFTNSVPKREGSNCIDYASSGNPEHLNLTYSILCTLLGLFLPLSVM 199		
DB	121 LAIWLVLELPLMTFTNSVPKREGSNCIDYASSGNPEHLNLTYSILCTLLGLFLPLSVM 180		
QY	200 CFYFKWVFLKRSQQOATLPLDKPQRLVLAIVFISILFTFTSMRDYLLMK 259		
DB	181 CFYFKWVFLKRSQQOATLPLDKPQRLVLAIVFISILFTFTSMRDYLLMK 240		
QY	260 PQ-GCTQKAISYITL 276		
DB	241 KQYQCTQVINSFYIVTR 258		
RESULT 10			
AA04389			
ID	AA04389 standard; Protein; 373 AA.		
XX	AA04389;		
XX	04-SEP-2001 (first entry)		
DT	Human P2-purinergic receptor subtype, P2Y1.		
DE	Human; P2-purinergic receptor; P2Y1; cardiant; vasotropic; thrombolytic;		
XX	cerebroprotective; gynaecological; ADP; adenosine 5'-diphosphate; angina;		
KW	myocardial infarction; ischaemic attack; preclampsia; bleeding disorder;		
KW	carotid endarterectomy; vascular graft surgery; brain disorder; migraine;		
KW	vascular injury; schizophrenia; eating disorder; depression; angioplasty;		
KW	peripheral vascular disease; platelet aggregation; restenosis; embolism;		
KW	thrombocytopenic purpura; stroke; pertussis toxin-sensitive G protein;		
XX	Gi; disseminated intravascular coagulation; thrombosis.		
XX	Homo sapiens.		

XX FH Key Location/Qualifiers
 FT Domain 53..76
 FT /label= Transmembrane_domain_1
 FT Domain 84..109
 FT /label= Transmembrane_domain_2
 FT Domain 124..148
 FT /label= Transmembrane_domain_3
 FT Domain 167..190
 FT /label= Transmembrane_domain_4
 FT Domain 217..241
 FT /label= Transmembrane_domain_5
 FT Domain 258..283
 FT /label= Transmembrane_domain_6
 FT Domain 304..329
 FT /label= Transmembrane_domain_7

XX WO200146454-A1.

XX 28-JUN-2001.

XX 26-DEC-2000; 2000WO-US34998.

XX 23-DEC-1999; 99US-0171622.

XX (CORT-) COR THERAPEUTICS INC.

XX Conley PB, Jantzen H, Ramakrishnan-Dubridge V, Julius DJ;

XX Hollopeter G;

XX WPI; 2001-418082/44.

XX Novel isolated ADP receptor, termed P2Y12 receptor polypeptide, useful for identifying binding partners and for diagnostic applications

XX Disclosure; Fig 5A; 108pp; English.

XX The invention relates to ADP (adenosine 5'-diphosphate) receptor, termed P2Y12 receptor and its corresponding cDNA molecule. P2Y12 receptor is the subtype of P2-purinergic receptor. The P2Y12 receptor is expressed selectively in the platelets and brain, and couples to a pertussis toxin-sensitive G protein (G_i). P2Y12 receptor is a G protein-coupled receptor that responds to ADP. The invention also relates to a method for identifying an agent which is useful for modulating acute myocardial infarction, unstable angina, chronic stable angina, transient ischaemic attacks, strokes, peripheral vascular disease, pre-eclampsia, deep venous thrombosis, embolism, disseminated intravascular coagulation, thrombotic thrombocytopenic purpura or a bleeding disorder; thrombotic and restenotic complications following angioplasty, carotid endarterectomy, post CABG (coronary artery bypass graft) surgery, vascular graft surgery, stent placements or insertion of endovascular devices and for P2Y12 receptor is useful for identifying binding partners and for diagnostic applications. P2Y12 receptor provides targets for screening synthetic small molecules and combinatorial or naturally occurring compound libraries to regulate platelet aggregation, vascular injury, or disease as well as schizophrenia, eating disorders, depression, migraine and other brain disorders. The present sequence is human P2-purinergic receptor subtype, P2Y1 related to the invention.

XX Sequence 373 AA;

XX Query Match 28.78; Score 474; DB 22; Length 373;

XX Best Local Similarity 37.38; Pred. No. 3.7e-39;

XX Matches 107; Conservative 55; Mismatches 117; Indels 8; Gaps 6;

XX 20 YLSAFVAIEIFGLGNVTVVGYLFCMKNNSSNYLFNLSTDSFAFLCTLPILKSY 79

XX 52 YLPVAVYLVFIIFGLGNVAIWVFMKPGSGISVYMFNLADELFLVTLPALIFY 111

XX 80 ANDKG-TYGDVLCISNRVLTNLTSLTFLTFISMDRYLLMKYPFHEFLQKEFALI 138

XX 112 FNKTDWIFGDAMCKLQRFIFVNLVYGLSILFLTCISAHRYSGVYVPLKSLGRKKNAICI 171

QY 139 SLAWALVTLEVLPMLTFFINS-VPKEEGSNCIDYASSGNPEHNLVLSLCTLLGLFIPLS 197
 Db 172 SVLVLLVVAISPILFYSGTGVRKNKTTICYDTSDEYLSRYFIYSMTTIVAMFCVPLV 231
 QY 198 VMCFYFKVWVFLKRRSQOQATAPL-DKQORLVLAUVIFSILETTPYHIMRNLRIASRL 256
 Db 232 LILGCGYGLIVRALIYKDLONS---PLRRKSIYLVIIIVFAVSYPHFVMTMNLRLARL 288
 QY 257 D-SWPGGCT-QKAIKSIYTLTREPFLNSAINPIFYFLMGDHYREML 301
 Db 289 DFQTPAMCAFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRRL 335

RESULT 11

AAU10983

ID AAU10983 standard; Protein; 373 AA.

XX AC AAU10983;

XX DT 12-MAR-2002 (first entry)

XX Purinergic receptor P2Y, G-protein coupled 1.

XX Purinergic receptor P2Y, G-protein coupled 1; P2RY1; anticoagulant; coagulant; platelet aggregation; haplotyping; drug screening; transgenic animal; human.

XX OS Homo sapiens.

XX WO200190117-A2.

XX 29-NOV-2001.

XX 21-MAY-2001; 2001WO-US16432.

XX 19-MAY-2000; 2000US-205996P.

XX (GENA-) GENAISSANCE PHARM INC.

XX Kazemi A, Koshiy B, Tanguay DA;

XX WPI; 2002-083074/11.

XX N-PSDB; AAS18599.

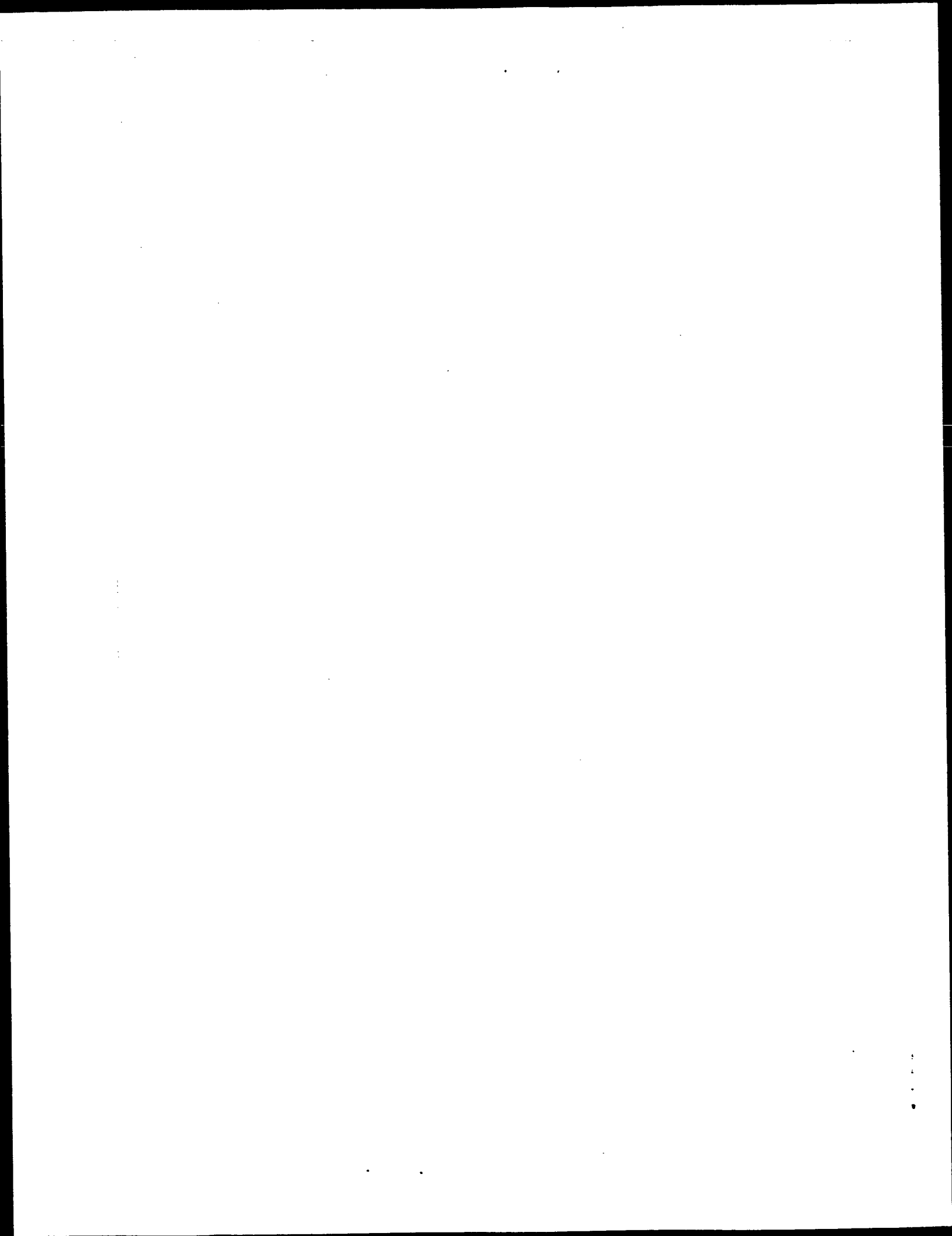
XX New purinergic receptor P2Y G-protein coupled 1 (P2RY1) gene polymorphic variants, useful e.g. in studying the expression and function of P2RY1 and screening candidate drugs for treating diseases related to P2RY1 activity

XX Claim 28; Fig 3; 79pp; English.

XX The invention relates to a novel isolated polypeptide comprising a sequence which is a polymorphic variant of a reference sequence for the purinergic receptor P2Y, G-protein coupled, 1 (P2RY1) protein or its fragment. The polymorphic variant comprises one or more variant amino acids selected from valine at a position 34 and glycine at a position 262. The polymorphic variants are useful in studying the expression and function of P2RY1, in expressing P2RY1 protein for use in screening for candidate drugs to treat diseases related to P2RY1 activity, in studying the effect of the variation on the biological activity of P2RY1, and the binding affinity of candidate drugs targeting P2RY1 for the treatment of disorders related to platelet aggregation. The haplotyping methods are useful in validating P2RY1 as a candidate target for treating a specific condition or disease predicted to be associated with P2RY1 activity, or in the design of clinical trials of candidate drugs for treating a specific condition or disease associated with P2RY1 activity. The transgenic animals are useful for studying expression of the P2RY1 isogenes in vivo, for in vivo screening and testing of drugs targeted against P2RY1 protein, and for testing the efficacy of therapeutic agents and compounds for disorders related to platelet aggregation in a biological system. The present sequence represents the amino acid sequence of human purinergic receptor P2Y, G-coupled protein 1 (P2RY1).

PF	21-MAY-2001; 2001WO-US16432.	
XX		
PR	19-MAY-2000; 2000US-205996P.	
XX		
PA	(GENA-) GENAISSANCE PHARM INC.	
XX		
PI	Kazemi A, Koshy B, Tanguay DA;	
XX		
DR	WPI; 2002-083074/11.	
XX		
XX	New purinergic receptor P2Y G-protein coupled 1 (P2RY1) gene	
PT	polymorphic variants, useful e.g. in studying the expression and	
PT	function of P2RY1 and screening candidate drugs for treating diseases	
PT	related to P2RY1 activity	
XX		
PS	Disclosure; Page -: 79pp; English.	
XX		
CC	The invention relates to a novel isolated polypeptide comprising a	
CC	sequence which is a polymorphic variant of a reference sequence for the	
CC	purinergic receptor P2Y, G-protein coupled, 1 (P2RY1) protein or its	
CC	fragment. The polymorphic variant comprises one or more variant amino	
CC	acids selected from valine at a position 34 and glycine at a position	
CC	262. The polymorphic variants are useful in studying the expression	
CC	and function of P2RY1, in expressing P2RY1 protein for use in screening	
CC	for candidate drugs to treat diseases related to P2RY1 activity, in	
CC	studying the effect of the variation on the biological activity of	
CC	P2RY1, and the binding affinity of candidate drugs targeting P2RY1 for	
CC	the treatment of disorders related to platelet aggregation. The	
CC	haplotyping methods are useful in validating P2RY1 as a candidate	
CC	target for treating a specific condition or disease predicted to be	
CC	associated with P2RY1 activity, or in the design of clinical trials of	
CC	candidate drugs for treating a specific condition or disease associated	
CC	with P2RY1 activity. The transgenic animals are useful for studying	
CC	expression of the P2RY1 isogenes in vivo, for in vivo screening and	
CC	testing of drugs targeted against P2RY1 protein, and for testing the	
CC	efficacy of therapeutic agents and compounds for disorders related to	
CC	platelet aggregation in a biological system. The present sequence	
CC	represents the amino acid sequence of human purinergic receptor P2Y,	
CC	G-coupled protein 1 (P2RY1), isoform #2.	
CC	Note: This sequence is not shown in the specification but is	
CC	derived from the human wild type P2RY1 sequence given in Figure 3 (see	
CC	AAU10983).	
XX		
SQ	Sequence 373 AA;	
	Query Match 28.3%; Score 467; DB 23; Length 373;	
	Best Local Similarity 36.9%; Pred. No. 1.9e-38;	
	Matches 106; Conservative 55; Mismatches 118; Indels 8; Gaps 6;	
Qy	20 YLSAFYAIEFIFGLGNVTWVFGYLFCKMKNSSNVYLFNLSISDFAFICTLPILIKSY 79	
Db	52 YLPAVYILVFIIGFLGNSVAIWMFVFKMPKPSGIVSMFNLAADFLYVLTLPALIFY 111	
Qy	80 ANDKG-TYGDVLCISNRYVLHTNLYTSILFLTFISMDRYLLMKYPFEHLQKKEAII 138	
Db	112 FNKTDWIFGDAMCKLQRFIEHVNLYGILFTCSIAHRYSGVYVPLKSLGRUKKNAICI 171	
Qy	139 SLAWALVTLVLPMLFINS-VPEKSGSCIDYVASSGNPEHNLIVSLCLTLGLGFILPLS 197	
Db	172 SVLVNLIIVVAISLPIFYSGTGVKRNKTYTCYDTSDEYLSRYFIYSMTTVMFCVPLV 231	
Qy	198 VMCFFYKMYVFLKRRSQOQATLPL-DKQORLVVLAVVIFSLFTPYHMRNRIASRL 256	
Db	232 LILGCGYLIVRALYVKDLDNS---PLRRKSYILGIIIVTFVAVSYIPIFVHMKTMNLRARL 288	
Qy	257 D-SWPGQCT-QKAIKSYITLFRPLAFINSAINPIFYLMGDHYREML 301	
Db	289 DFQTPAMCAFNDRVATYQVTRGLASLNSCDPILILFLAGDTFRRL 335	
	RESULT 14	
	AAU04375	
ID	AAU04375 standard; Protein; 337 AA.	

XX	AAU04375;	
AC		
XX	23-OCT-2001 (first entry)	
DT		
XX		
DE	Human G-protein coupled receptor, hrUP21.	
XX		
XX	Human; G-protein coupled receptor; GPCR; hrUP21; agonist;	
KW	inverse agonist; lung cancer.	
KW		
XX	Homo sapiens.	
OS		
XX	WO200136471-A2.	
XX		
PN		
XX	25-MAY-2001.	
XX		
PD		
XX	16-NOV-2000; 2000WO-US31509.	
PF		
XX	17-NOV-1999; 99US-0166088.	
XX	17-NOV-1999; 99US-0166099.	
PR	17-NOV-1999; 99US-0166369.	
PR	23-DEC-1999; 99US-0171900.	
PR	23-DEC-1999; 99US-0171901.	
PR	23-DEC-1999; 99US-0171902.	
PR	23-DEC-1999; 2000US-0181749.	
PR	11-FEB-2000; 2000US-0189258.	
PR	14-MAR-2000; 2000US-0189259.	
PR	10-APR-2000; 2000US-0195898.	
PR	10-APR-2000; 2000US-0195899.	
PR	10-APR-2000; 2000US-0196078.	
PR	28-APR-2000; 2000US-0200419.	
PR	12-MAY-2000; 2000US-0203630.	
PR	12-JUN-2000; 2000US-0210741.	
PR	12-JUN-2000; 2000US-0210982.	
PR	21-AUG-2000; 2000US-0226760.	
PR	26-SEP-2000; 2000US-0235418.	
PR	26-SEP-2000; 2000US-0235779.	
PR	20-OCT-2000; 2000US-0242332.	
PR	20-OCT-2000; 2000US-0242343.	
XX	(AREN-) ARENA PHARM INC.	
PA		
XX	Chen R, Dang HT, Lowitz KP;	
PI		
XX	WPI; 2001-355616/37.	
DR	N-P5DB; AAS07948.	
DR		
XX	Endogenous and non-endogenous versions of human G-protein coupled	
PT	receptors for direct identification of candidate compounds as agonists,	
PT	inverse agonists or partial agonists for use as therapeutic agents -	
XX	Claim 53; Page 114-115; 160pp; English.	
PS		
XX	The sequence represents a human G-protein coupled receptor (GPCR),	
CC	hrUP21. The endogenous and non-endogenous, constitutively activated	
CC	versions of human G-protein coupled receptors (GPCR), are useful for	
CC	direct identification of candidate compounds as receptor agonists,	
CC	inverse agonists or partial agonists having applicability as therapeutic	
CC	agents for treating diseases related to GPCR, e.g. lung cancer.	
CC	Non-endogenous version of human GPCRs are also utilized in research	
CC	settings and in vitro and in vivo system, incorporating GPCRs can be	
CC	utilised to elucidate and understand the roles these receptors	
CC	play in the human condition, both normal and diseased.	
XX	Sequence 337 AA;	
SQ		
	Query Match 25.8%; Score 426; DB 22; Length 337;	
	Best Local Similarity 34.4%; Pred. No. 2.2e-34;	
	Matches 101; Conservative 57; Mismatches 126; Indels 10; Gaps 6;	
Qy	17 LNKYYLSAFYAIEFIFGLGNVTWVFGYLFCKMKNSSNVYLFNLSISDFAFICTLPILI 76	
Db	31 LKMHLPVYIGIIFGLVPGNAVVIITYIFKMRPKWSSTIIMLNLACTIONLLYITSLPFLI 90	



Db 65 LSVSOLAFLCTLPMLIRSYANGNWIYGVLCISNRVYVLANLYTSILFTFISTDRYLI 124
QY 121 KYPFREHFLQKKEFAILLISLAWALVTLEVLPLMTFTFINSVPKKEGSCNDYASSGNPEHN 180
Db 125 KYPFREHLLQKKEFAILLISLAWALVTLEVLPLMTFTFINSVPKKEGSCNDYASSGNPEHN 184
QY 181 LIYSCLTLLGLFLPLSVCMCFYKVMVFLKRRSQOQATPLDKPQRLVLAUVIFSIL 240
Db 185 LIYSCLTLLGLFLPLSVCMCFYKVMVFLKRRSQOQATPLDKPQRLVLAUVIFSIL 244
QY 241 FPPYHIMRNLRATSLDSDWPO-GCTQKATKSIYTLTRPLAFNSAINPIFYFLMGDHYRE 299
Db 245 FPPYHIMRNLRATSLDSDWPO-GCTQKATKSIYTLTRPLAFNSAINPIFYFLMGDHYRE 304
QY 300 MLISKRFQYFKSLTSF 315
Db 305 MLMNQLRHNFKSLTSF 320

RESULT 2
US-08-749-707-2
: Sequence 2, Application US/08749707
: Patent No. 6063582
: GENERAL INFORMATION:
: APPLICANT: Conley, Pamela B.
: APPLICANT: Jantzen, Hans-Michael
: TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
: STREET: 1800 M Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20036-5869
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/749,707
: FILING DATE: 15-NOV-1996
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Adler, Reid G.
: REGISTRATION NUMBER: 30,988
: REFERENCE/DOCKET NUMBER: 044481-5010-01-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-467-7000
: TELEFAX: 202-467-7176
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 334 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-749-707-2

Query Match 74.3%; Score 1226.5; DB 3; Length 334;
Best Local Similarity 71.5%; Pred. No. 3.5e-93;
Matches 226; Conservative 42; Mismatches 47; Indels 1; Gaps 1;
QY 1 MAONLSCEWLAATEAILNKYLSAFYAEFTFGLGNVTGVVFGYLFCKMKNSSNVYLFN 60
Db 5 MAWNAATKNWLAEEALKEKYLISFYEGFEVGVGLNVTGVVFGYLFCKMKNSSNVYLFN 64
QY 61 LSISDAFLCTLPILIKSYANDKGYGVGVLCISNRVYVLANLYTSILFTFISMDRLLM 120
Db 65 LSVSOLAFLCTLPMLIRSYANGNWIYGVLCISNRVYVLANLYTSILFTFISTDRYLI 124
QY 121 KYPFREHFLQKKEFAILLISLAWALVTLEVLPLMTFTFINSVPKKEGSCNDYASSGNPEHN 180

Db 125 KYPFREHLLQKKEFAILLISLAWALVTLEVLPLMTFTFINSVPKKEGSCNDYASSGNPEHN 184
QY 181 LIYSCLTLLGLFLPLSVCMCFYKVMVFLKRRSQOQATPLDKPQRLVLAUVIFSIL 240
Db 185 LIYSCLTLLGLFLPLSVCMCFYKVMVFLKRRSQOQATPLDKPQRLVLAUVIFSIL 244
QY 241 FPPYHIMRNLRATSLDSDWPO-GCTQKATKSIYTLTRPLAFNSAINPIFYFLMGDHYRE 299
Db 245 FPPYHIMRNLRATSLDSDWPO-GCTQKATKSIYTLTRPLAFNSAINPIFYFLMGDHYRE 304
QY 300 MLISKRFQYFKSLTSF 315
Db 305 MLMNQLRHNFKSLTSF 320

RESULT 3
US-08-559-524A-4
: Sequence 4, Application US/08559524A
: Patent No. 5871963
: GENERAL INFORMATION:
: APPLICANT: Conley, Pamela B.
: APPLICANT: Jantzen, Hans-Michael
: TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
: STREET: 1800 M Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20036-5869
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/559,524A
: FILING DATE: 15-NOV-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Adler, Reid G.
: REGISTRATION NUMBER: 30,988
: REFERENCE/DOCKET NUMBER: 044481-5010-00-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-467-7000
: TELEFAX: 202-467-7176
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 373 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-559-524A-4

Query Match 28.7%; Score 474; DB 2; Length 373;
Best Local Similarity 36.6%; Pred. No. 1.6e-31;
Matches 105; Conservative 58; Mismatches 116; Indels 8; Gaps 6;
QY 20 YLSAFYAEFTFGLGNVTGVVFGYLFCKMKNSSNVYLFNLSISDAFLCTLPILIKSY 79
Db 52 YLPAVAVILVFIIGFLGNSVAIWMFVHMKPWSGISVYMFENALADFLVLTLPALIFY 111
QY 80 ANDKG-TYGDVLCISNRVYVLANLYTSILFTFISMDRLLMKYPREHFLQKKEFAILI 138
Db 112 FNKTDFGDMCKLQRFIFHVNLYGSLFLTCISAHRYSGVVYPLKSLGRLKKNNAVYI 171
QY 139 SLAWALVTLEVLPLMTFTFINS-VPKKEGSCNDYASSGNPEHNLYTSILCTLLGLFLPLS 197
Db 172 SVLWLVVVGISPLIFYSGTIGIRKNKTKITCYDTTSDVYLSRYFYISMCCTVAMFCVPLV 231
QY 198 VMCFYKVMVFLKRRSQOQATPL-DKPQRLVLAUVIFSILFTFYPYHIMRNLRATSL 256

Db 232 LILGCGYGLIVRALIYKDLNDS---PLRKSIIYLIIVLTVFAVSIYIPFHWKTMNLRARL 288
 QY 257 D-SWPGCT-QKATKSIYTLTRPLAFNSAINPIFYFLMGDHYREML 301
 Db 289 DFQTPMCAFNDRYATYQVTRGLASLNSCVDPILYFLAGDTFRRL 335

RESULT 4

US-08-749-707-4
 ; Sequence 4, Application US/08749707
 ; Patent No. 6063582
 ; GENERAL INFORMATION:
 ; APPLICANT: Conley, Pamela B.
 ; APPLICANT: Jantzen, Hans-Michael
 ; TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
 ; STREET: 1800 M Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20036-5869
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/749,707
 ; FILING DATE: 15-NOV-1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Adler, Reid G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 044481-5010-01-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-467-7000
 ; TELEFAX: 202-467-7176
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 373 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-749-707-4

Query Match 28.7%; Score 474; DB 3; Length 373;
 Best Local Similarity 36.6%; Pred. No. 1.6e-31;
 Matches 105; Conservative 58; Mismatches 116; Indels 8; Gaps 6;
 QY 20 YLSAFYALFTFGLGNVTVVGYLFCMKNNNSNYLNLNLSISDEAFGLCTLPILIKSY 79
 Db 52 YVLPVAVYILVFIIGLGNLSVAVMFWFKPWSGISVYFNLALADFLYVITFLPALIPIY 111
 QY 80 ANDKG-TYGDVLCISNRVYHNLVTSILFTFTISMDRYLLMKYPFRHFQKKEFALLI 138
 Db 112 FNKTWIFGDAMCKLQRFIFHVNYSILFTLCISAHRYSGVYVYPLKSLGRLLKKNVYI 171
 QY 139 SLAVVALVTVLEVLPLMTFINS-VPEEGSCNIDYASSNPENHLYISCLLLGLFLPLS 197
 Db 172 SVLWLVVVGISPLIFSYGGIRKNKTTICYDTSDEYLSRYFIYSMCTTVAMFCVPLV 231
 QY 198 VMCFFYKVMVFLKRRSQOQATAPL-DKQRLVVLAVVIFSLFTPTVHMRNLRISRL 256
 Db 232 LILGCGYGLIVRALIYKDLNDS---PLRKSIIYLIIVLTVFAVSIYIPFHWKTMNLRARL 288
 QY 257 D-SWPGCT-QKATKSIYTLTRPLAFNSAINPIFYFLMGDHYREML 301
 Db 289 DFQTPMCAFNDRYATYQVTRGLASLNSCVDPILYFLAGDTFRRL 335

RESULT 5

US-08-513-974B-374
 ; Sequence 374, Application US/08513974B
 ; Patent No. 6114139
 ; GENERAL INFORMATION:
 ; APPLICANT: Hinuma, Shuji
 ; APPLICANT: Hosoya, Masaki
 ; APPLICANT: Fujii, Ryo
 ; APPLICANT: Ohtaki, Tetsuya
 ; APPLICANT: Fukusumi, Shoji
 ; APPLICANT: Ohgi, Kazuhiro
 ; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
 ; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
 ; NUMBER OF SEQUENCES: 380
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 ; STREET: 130 Water Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/513,974B
 ; FILING DATE: 14-SEP-1995
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/JF95/01599
 ; FILING DATE: 10-AUG-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 7-093989
 ; FILING DATE: 19-AUG-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 7-057186
 ; FILING DATE: 16-MAR-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 7-007177
 ; FILING DATE: 20-JAN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-326611
 ; FILING DATE: 28-DEC-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-270017
 ; FILING DATE: 02-NOV-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-236357
 ; FILING DATE: 30-SEP-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-236356
 ; FILING DATE: 30-SEP-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-189274
 ; FILING DATE: 11-AUG-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-189273
 ; FILING DATE: 11-AUG-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-189272
 ; FILING DATE: 11-AUG-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Resnick, David S.
 ; REGISTRATION NUMBER: 34,235
 ; REFERENCE/DOCKET NUMBER: 45753
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-523-3400
 ; TELEFAX: 617-523-6440
 ; INFORMATION FOR SEQ ID NO: 374:
 ; SEQUENCE CHARACTERISTICS:

us-09-891-138a-2.rai

Thu Apr 3 10:42:34 2003

LENGTH: 362 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-513-974B-374

Query Match 28.7%; Score 473.5; DB 3; Length 362;
Best Local Similarity 35.9%; Pred. No. 1.7e-31;
Matches 110; Conservative 57; Mismatches 122; Indels 17; Gaps 8;

QY 10 WLA-----TEAILNK-----YILSAFVAIEFIFGLLGNVTTFVFGYLCFCKMKNSSNVYLFN 60
DB 22 WAAGNATTKSLTKTGQFYLLPTVYILVITGFLGNSVAIWMFVHMRPWSGISVYMFN 81
QY 61 LSTDSFAFLCTLPILIKSYANDKG--TYGDVLCISNRYVLTNLYSILFLFISMDRYLL 119
DB 82 LALADFLYVLTLPALIFYFENKTDWIFGDMCKLQRFIFHVNLYSILFLTCISVHRVTG 141
QY 120 MKYPFHEFLQKEFAILSLAVWALVTLEVLPLMTFTIN-SVPKEEGSNCIDYASSGNPE 178
DB 142 VVHPLKSLGRKKNAVYSSSLWALVAVIAPIFYSGIGYVRNKTYTCYDTTADAYLR 201
QY 179 HNLISLCLTLGLFLPLSLMCFYFYKMYVFLKRRSQOQATLPL-DKQRLVLAIVIF 237
DB 202 SYFVSMCTVFECFPIFVILGCGYGLIVKALYKDLDS---PLRRKSTYLVIIIVLTFV 258
QY 238 SILFTYHMRNRIASRLD-SWPGCT-QKAISYITLTPLATLNSAINPIFYFLMGD 295
DB 259 AVSYLPFHVMTLNLRALDFTQPMCAFNDKYATYQVTRGLASLNSCVDPILYFLAGD 318
QY 296 HYREML 301
DB 319 TFRRL 324

RESULT 6

US-08-513-974B-373
Sequence 373, Application US/08513974B
Patent No. 6114139

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Ohtaki, Tetsuya
APPLICANT: Fukusumi, Shoji
APPLICANT: Ohgi, Kazuhiro
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
NUMBER OF SEQUENCES: 380
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513,974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-326611
FILING DATE: 28-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236356
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1945
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 373:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-513-974B-373

Query Match 22.5%; Score 370.5; DB 3; Length 373;

Best Local Similarity 33.5%; Pred. No. 4.6e-23;

Matches 94; Conservative 50; Mismatches 130; Indels 7; Gaps 5;

QY 19 KY-YLSAFYAIEFIFGLLGNVTTFVFGYLCFCKMKNSSNVYLFNLSISDFALCTLPILIK 77
DB 32 KYVLLPVSYGVCVGLGCLNVVALYIFLCRLKWNASTYMFHLAVSDLSYAAASLPLVY 91
QY 78 SYA-NDKGTGYDVLICISNRYVLTNLYTSILFLFISMDRYLLMKYPPREHFLQKEFAI 136
DB 92 YYAGDHWPFFSTVLCVKLVRFYLTNLYCSILFLTCISVHRCLGVLRPLHSLRWGRARYAR 151
QY 137 LISLAWALVTLEVLPLMTFTINSPKEEGSNCIDYASSGNPEHNLISLCTLLGLFLIPL 196
DB 152 RVAAVVWLVLACQAPVLYFVTTSVRGTRITCHDTSARELFSHFVAYSSVLMGLLFAVPP 211
QY 197 SVMCFEYKMYVFLKRRSQOQATLPLDKPQ--RLVLAIVVIFSLFTPTPHMNRLIAS 254
DB 212 SVILVCYVLMARRLLKPAYGTGGPLPRKRSVTRIALVLAFLCFLPFHVRTLYYSF 271
QY 255 RLDSWPGQ--TOKAISKIYTLTRPLAFNSAINPIFYFLMG 294
DB 272 R--SLDLSCHTLNAINMAYKITRPLASANSCLDPVLYFLAG 310

RESULT 7

US-08-153-848-28
Sequence 28, Application US/08153848
Patent No. 5759804

GENERAL INFORMATION:

APPLICANT: Godiska, Ronald

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ADDRESSEE: Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/299,843A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 09/088,337
 FILING DATE: 01-JUN-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/153,848
 FILING DATE: 17-NOV-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/977,452
 FILING DATE: 17-NOV-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Jill E. Uhl
 REGISTRATION NUMBER: 43,213
 REFERENCE/DOCKET NUMBER: 27866/32059B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX:
 INFORMATION FOR SEQ ID NO: 28:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 355 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-299-843A-28

Query Match 21.4%; Score 353; DB 3; Length 355;
 Best Local Similarity 28.1%; Pred. No. 1.2e-21;
 Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;
 QY 16 ILNKYLSAFAIEIFGLGNVTWVFGYLCMKNNSSNVYLFNLSISDFAPLCTLPIL 75
 Db 28 VFGTVLSIFSVIFAIGLVGNLLVVFALTNKKPKSVTDIYLLNLSLDFVATLPFW 87
 QY 76 IKSYANDKGTGYGVLCISNRYVLTNLTYSILFTFISMDRYLLMKYPFHEFLQKKEA 135
 Db 88 THYLINEKGLH-NAMCKFTTAFFFIGFGSIFFTVISIDRYLAIVLAANSNNRTVOHG 146
 QY 136 ILISLAWALVTLEVPMLTFINSVPKKEGNCI-DYASSGNPEHNLISLCTLLGLFLI 194
 Db 147 VTISLGWAAAILVAAPQPMF---TKOKENECGDYPEVLQEIWPLRVNVTNFGFL 202
 QY 195 PLSVMCFYFKMV-VFLKRRSQOQATALPLDKPQLRVLAIVFISILFTPYHIMNLR 253
 Db 203 PLLIMSYCFRIQTLSCKNHKKAKI-----KLILLVIVVFFLFTWPTYNMIFLET 256
 QY 254 SRLDSWPGCTOKAISKIYTLRPLAFNSAINPIFFELMGDHYREMLISKFRQYFKSL 312
 Db 257 KLYDFPSCDMRKDLRLALSVTETVAFSHCCLNPLIYAFAGEKFRYL---YHLYGKCL 312

RESULT 10
 US-09-299-843A-32
 ; Sequence 32, Application US/09299843A
 ; Patent No. 6107475
 ; GENERAL INFORMATION:
 ; APPLICANT: Godiska, Ronald

APPLICANT: Gray, Patrick W.
 APPLICANT: Schweikart, Vicki L.
 TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors
 NUMBER OF SEQUENCES: 66
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ADDRESSEE: Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/299,843A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 09/088,337
 FILING DATE: 01-JUN-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/153,848
 FILING DATE: 17-NOV-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/977,452
 FILING DATE: 17-NOV-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Jill E. Uhl
 REGISTRATION NUMBER: 43,213
 REFERENCE/DOCKET NUMBER: 27866/32059B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX:
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 355 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-299-843A-32

Query Match 21.4%; Score 353; DB 3; Length 355;
 Best Local Similarity 28.1%; Pred. No. 1.2e-21;
 Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;
 QY 16 ILNKYLSAFAIEIFGLGNVTWVFGYLCMKNNSSNVYLFNLSISDFAPLCTLPIL 75
 Db 28 VFGTVLSIFSVIFAIGLVGNLLVVFALTNKKPKSVTDIYLLNLSLDFVATLPFW 87
 QY 76 IKSYANDKGTGYGVLCISNRYVLTNLTYSILFTFISMDRYLLMKYPFHEFLQKKEA 135
 Db 88 THYLINEKGLH-NAMCKFTTAFFFIGFGSIFFTVISIDRYLAIVLAANSNNRTVOHG 146
 QY 136 ILISLAWALVTLEVPMLTFINSVPKKEGNCI-DYASSGNPEHNLISLCTLLGLFLI 194
 Db 147 VTISLGWAAAILVAAPQPMF---TKOKENECGDYPEVLQEIWPLRVNVTNFGFL 202
 QY 195 PLSVMCFYFKMV-VFLKRRSQOQATALPLDKPQLRVLAIVFISILFTPYHIMNLR 253
 Db 203 PLLIMSYCFRIQTLSCKNHKKAKI-----KLILLVIVVFFLFTWPTYNMIFLET 256
 QY 254 SRLDSWPGCTOKAISKIYTLRPLAFNSAINPIFFELMGDHYREMLISKFRQYFKSL 312
 Db 257 KLYDFPSCDMRKDLRLALSVTETVAFSHCCLNPLIYAFAGEKFRYL---YHLYGKCL 312

RESULT 11
 US-09-088-337B-28

; Sequence 28, Application US/09088337B
; Patent No. 6348574

GENERAL INFORMATION

APPLICANT: CODING

AFFILIATION: GODISKA, RONALD
Grav. Dates: 1943-1944

Gray, Patrick W.

Schweikart, Vicki L.

TITLE OF INVENTION: No. 6348574e1 Seven Transmembrane Receptors

NUMBER OF SEQUENCES: 66

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Associates, Inc.

BOBIN

STREET: 6300 Sears Tower 222 South Wacker Drive
 Detail

STREET: 6300 SE
CITY: Chicago

CITY: Chicago
STATE: Illinois

STATE: Illinois

COUNTRY: US

ZIP: 60606

TER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS

OPERATING SYSTEM: PC
SOFTWARE: DATETYPE D

SOFTWARE: PatentIn Release #1.0, Ver
 MIT ADDITIONAL FILES

NT APPLICATION DATA:

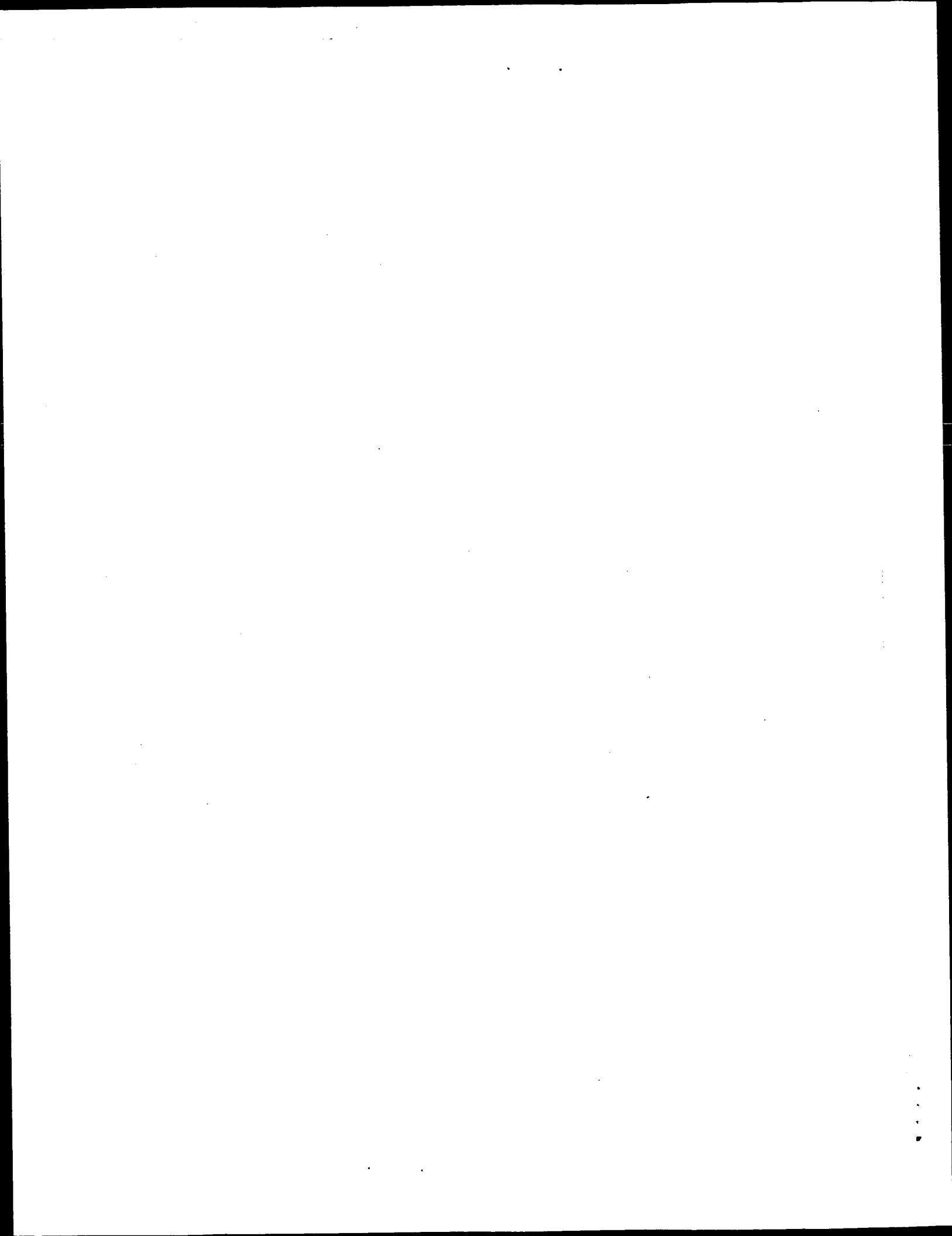
RESULT 15
US-08-513-974B-39
; Sequence 39, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji

RESULT 14
PCT-US93-11153-32
: Sequence 32. Application PC/TUS9311153

APPLICANT: Hosoya, Masaki
 APPLICANT: Fujii, Ryo
 APPLICANT: Ohtaki, Tetsuya
 APPLICANT: Fukusumi, Shoji
 APPLICANT: Ohgi, Kazuhiro
 TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
 TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
 NUMBER OF SEQUENCES: 380
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 STREET: 130 Water Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/513,974B
 FILING DATE: 14-SEP-1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP95/01599
 FILING DATE: 10-AUG-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 7-093989
 FILING DATE: 19-AUG-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 7-057186
 FILING DATE: 16-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 7-007177
 FILING DATE: 20-JAN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-326611
 FILING DATE: 28-DEC-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-270017
 FILING DATE: 02-NOV-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-236357
 FILING DATE: 30-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-236356
 FILING DATE: 30-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-189274
 FILING DATE: 11-AUG-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-189273
 FILING DATE: 11-AUG-1945
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-189272
 FILING DATE: 11-AUG-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Resnick, David S.
 REGISTRATION NUMBER: 34,235
 REFERENCE/DOCKET NUMBER: 45753
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-523-3400
 TELEFAX: 617-523-6440
 INFORMATION FOR SEQ ID NO: 39:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 328 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide

Query Match 20.5%; Score 338; DB 3; Length 328;
 Best Local Similarity 29.6%; Pred. No. 1.8e-20;
 Matches 84; Conservative 54; Mismatches 136; Indels 10; Gaps 4;
 QY 22 LSAYATIEFIFGLGNVTWVFGYLFMCKNWSNVVLENLSTSDFAFLCTLPILIKSYA- 80
 Db 29 LTPVSVLVVGLPLNICVIAQICASRRTLTRSAVYTNLALADLMYACSLPLLIYNAR 88
 QY 81 NDKGTGVDVLCISNRYVLTNLYTSLFTFTSMRDYLLMKYPFRE-HFLOKKEFAILIS 139
 Db 89 GDHWPEGLACRFVRFELFYANLHGSILFTLCISFQRYLGICHPLASWIKRGRRAAVVC 148
 QY 140 LAWVALVTLEVLPLMTTFINSVPKEGSCIDYASSGNPEHNLIIYSLCTLTLLGLPLISVM 199
 Db 149 GVVWLVAVTAQCLPTAVFAATGIQRNRTVCYDLSPPILSTRYLPYGMALTVIGELLPTAL 208
 QY 200 CFFYYKVVVFLKRRSQQTALPL-----DKPQLVWLAVVIFSLFPPYHIMNLRAS 254
 Db 209 LACYCRMA---RRLCRQDGPAGPVAQERRSKAARMVVAAVFAISFLPFHTTAYLAV 265
 QY 255 RLDSWPGQCTOKAISKIYTLTRPLAFNLAINPIFYFLMGDHYR 298
 Db 266 RSTPGVSCPVLTEPFAAAAYKTRPFASVNSVLDPLILFYFTQOKFR 309

Search completed: April 2, 2003, 15:18:31
 Job time : 20 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 15:17:12 ; Search time 23 Seconds
(without alignments)
842.614 Million cell updates/sec

Title: US-09-891-138A-2
Perfect score: 1650
Sequence: 1 MAQNLSCEWLATEAILNKY.....REMLISKFROYFKSILTSFRT 317

Scoring table: BLOSUM62

Searched: 248812 seqs, 61136040 residues
Total number of hits satisfying chosen parameters: 248812

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 s
```

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/us08_NEW_PUB.pcp.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	1231.5	74.6	334	10	US-09-765-034-2	Sequence 2, Appli
2	1234.5	74.2	334	9	US-10-270-587-2	Sequence 2, Appli
3	479	29.0	373	9	US-10-092-135-7	Sequence 7, Appli
4	475	28.8	299	9	US-10-270-144-4	Sequence 4, Appli
5	474	28.7	373	9	US-10-092-135-5	Sequence 5, Appli
6	474	28.7	373	9	US-10-092-135-6	Sequence 6, Appli
7	472.5	28.6	362	9	US-10-092-135-3	Sequence 3, Appli
8	472.5	28.6	362	9	US-10-092-135-4	Sequence 4, Appli
9	426	25.8	337	9	US-10-023-775B-2	Sequence 2, Appli
10	426	25.8	337	9	US-10-270-144-2	Sequence 2, Appli
11	426	25.8	337	10	US-09-943-798-4	Sequence 2, Appli
12	373.5	22.6	341	9	US-09-570-587-3	Sequence 3, Appli
13	366	22.2	276	10	US-09-943-798-2	Sequence 2, Appli
14	353	21.4	355	10	US-09-789-482-4	Sequence 4, Appli
15	353	21.4	355	10	US-09-789-486-4	Sequence 4, Appli
16	336.5	20.4	360	9	US-10-120-394-20	Sequence 20, Appli
17	336.5	20.4	360	9	US-09-764-413-20	Sequence 20, Appli
18	336.5	20.4	360	10	US-09-837-446-2	Sequence 2, Appli
19	336.5	20.4	360	10	US-09-796-744-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1

US-09-765-034-2
: Sequence 2, Application US/09765034

```

: Patent No. US20020137887A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Hedrick, Joseph A.
: APPLICANT: Lachowicz, Jean E.
: APPLICANT: Wang, Wei
: APPLICANT: Gustafson, Eric L.
:
: TITLE OF INVENTION: Adenosine Receptor
:
: FILE REFERENCE: CN01084
: CURRENT APPLICATION NUMBER: US/09/765,034
: CURRENT FILING DATE: 2001-01-17
: NUMBER OF SEQ ID NOS: 4
:
: SOFTWARE: Patent In Ver. 2.1
:
: SEQ ID NO 2
:
: LENGTH: 334
:
: TYPE: PRT
:
: ORGANISM: homo sapiens
:
: US-09-765-034-2

```

[illegible]

QY	20	YYLSAFYAIEFIFGLGNVTVVGYLFCMKNNSSNYLFNLSISDPAFLCTLPILIKSY	79
Db	11	YYLPAVYILVFIIGELGSAIWMFVPHKPSWISGYMFENALADFLVYLLVTPALIFY	70
QY	80	ANDKG-TYGDVLCISNRVVLHNLNLTSTLFTFTISMDRYLLMKYPREHFLQKFAILI	138
Db	71	FNKTDWIFGDAMCKLQRFIFHNLVYGSILFTCTCSAHRYSGVVPYPLKSLGRLLKKNAYV	130
QY	139	SLAVNALVTLEVPMLTFIN-SVPKEEGSNCIDYASSGNPEHNHILYSLCITLLGFLPI	197
Db	131	SVLWMLVIVVAISPILFYSGGTGRKNKTVTCITDTSNDLYRSFYISMCYTVAMFCIPL	189

QY 198 VMCFYYKMWVFLKRRSQOATAPL-DKPORLVVLAIVFISILFTPTPHIMNRLRIASRL 256
 Db 190 VLIIGCGVGLV--KALIYNDLNSPLRRKSIYLVIIIVTFVFAVSYPFHVMKTMNLRARL 247
 QY 257 D-SWPGCT-OKAISIYTLTRPLAFNSAINPIFYFLMGDHYREML 301
 Db 248 DFQTPMCAFNDRVYATQVTRGLASLNSCVDPILYFLAGDTFRRL 294

RESULT 5
 US-10-092-135-5
 ; Sequence 5, Application US/10092135
 ; Publication No. US20030054374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
 ; FILE REFERENCE: D0134.NP
 ; CURRENT APPLICATION NUMBER: US/10/092,135
 ; PRIOR FILING DATE: 2002-03-06
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR FILING DATE: 2001-03-07
 ; NUMBER OF SEQ ID NOS: 75
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 373
 ; TYPE: PRT
 ; ORGANISM: BOS TAURUS
 US-10-092-135-5

Query Match 28.7%; Score 474; DB 9; Length 373;
 Best Local Similarity 36.6%; Pred. No. 2.9e-34;
 Matches 105; Conservative 58; Mismatches 116; Indels 8; Gaps 6;
 QY 20 YLSAFAYAEFFIPLGLGNVTVVFGVLFCKMKNSSNVYLFNLSISDFAFCLTLPILIKSY 79
 Db 52 YLPAVYILVFIIGLGNVTVVFGVLFCKMKNSSNVYLFNLSISDFAFCLTLPILIKSY 111
 QY 80 ANDKG--TYGDVLCISNRVYVHTNLYTSILFTFISMDRYLLMKYPFREHFLQKKEFALLI 138
 Db 112 FNKTDWIFGDAMCKLQRFIFHVNLYGSLFTCISAHRYSGVVPYPLKSLGRKKNAICI 171
 QY 139 SLAVWALVTLEVPMLTFINS--VPKEGSGNCIDYASSGNPEHNLISLCLTLGLFLIPLS 197
 Db 172 SVLWMLIVVVAISPILFYSGTGVKRNKTIICYDTTSDYLSRYFIYSMCTTVAMFCVPLV 231
 QY 198 VMCFYYKMWVFLKRRSQOATAPL-DKPORLVVLAIVFISILFTPTPHIMNRLRIASRL 256
 Db 232 LILGCGVGLVRAIYKDLNDS---PLRRKSYLVIIIVTFVFAVSYPFHVMKTMNLRARL 288
 QY 257 D-SWPGCT-OKAISIYTLTRPLAFNSAINPIFYFLMGDHYREML 301
 Db 289 DFQTPMCAFNDRVYATQVTRGLASLNSCVDPILYFLAGDTFRRL 335

RESULT 6
 US-10-092-135-6
 ; Sequence 6, Application US/10092135
 ; Publication No. US20030054374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
 ; FILE REFERENCE: D0134.NP
 ; CURRENT APPLICATION NUMBER: US/10/092,135
 ; PRIOR FILING DATE: 2002-03-06
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR FILING DATE: 2001-03-07
 ; NUMBER OF SEQ ID NOS: 75
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: GALLUS GALLUS
 US-10-092-135-6

; NUMBER OF SEQ ID NOS: 75
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 373
 ; TYPE: PRT
 ; ORGANISM: HOMO SAPIENS
 US-10-092-135-6
 Query Match 28.7%; Score 474; DB 9; Length 373;
 Best Local Similarity 37.3%; Pred. No. 2.9e-34;
 Matches 107; Conservative 55; Mismatches 117; Indels 8; Gaps 6;
 QY 20 YLSAFAYAEFFIPLGLGNVTVVFGVLFCKMKNSSNVYLFNLSISDFAFCLTLPILIKSY 79
 Db 52 YLPAVYILVFIIGLGNVTVVFGVLFCKMKNSSNVYLFNLSISDFAFCLTLPILIKSY 111
 QY 80 ANDKG--TYGDVLCISNRVYVHTNLYTSILFTFISMDRYLLMKYPFREHFLQKKEFALLI 138
 Db 112 FNKTDWIFGDAMCKLQRFIFHVNLYGSLFTCISAHRYSGVVPYPLKSLGRKKNAICI 171
 QY 139 SLAVWALVTLEVPMLTFINS--VPKEGSGNCIDYASSGNPEHNLISLCLTLGLFLIPLS 197
 Db 172 SVLWMLIVVVAISPILFYSGTGVKRNKTIICYDTTSDYLSRYFIYSMCTTVAMFCVPLV 231
 QY 198 VMCFYYKMWVFLKRRSQOATAPL-DKPORLVVLAIVFISILFTPTPHIMNRLRIASRL 256
 Db 232 LILGCGVGLVRAIYKDLNDS---PLRRKSYLVIIIVTFVFAVSYPFHVMKTMNLRARL 288
 QY 257 D-SWPGCT-OKAISIYTLTRPLAFNSAINPIFYFLMGDHYREML 301
 Db 289 DFQTPMCAFNDRVYATQVTRGLASLNSCVDPILYFLAGDTFRRL 335

RESULT 7
 US-10-092-135-3
 ; Sequence 3, Application US/10092135
 ; Publication No. US20030054374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
 ; FILE REFERENCE: D0134.NP
 ; CURRENT APPLICATION NUMBER: US/10/092,135
 ; PRIOR FILING DATE: 2002-03-06
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR FILING DATE: 2001-03-07
 ; NUMBER OF SEQ ID NOS: 75
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: GALLUS GALLUS
 US-10-092-135-3

Query Match 28.6%; Score 472.5; DB 9; Length 362;
 Best Local Similarity 35.9%; Pred. No. 3.8e-34;
 Matches 110; Conservative 58; Mismatches 121; Indels 17; Gaps 8;
 QY 10 WLA-----TEAILNK-----YYLSAFAYAEFFIPLGLGNVTVVFGVLFCKMKNSSNVYLFN 60
 Db 22 WAAGNATTKSLTKTGFQFYVLPVYILVFIITGFLGNSVAIMVFMVFMRPWSGISVYMEN 81
 QY 61 LSISDFAFCLTLPILIKSYANDKG--TYGDVLCISNRVYVHTNLYTSILFTFISMDRYLL 119
 Db 82 LALADELYVLTLPALIFFYFNKTDWIFGDVYMKLQRFIFHVNLYGSLFTCISVHRVTG 141
 QY 120 MKYPFREHFLQKKEFALLISLAVWALVTLEVPMLTFINS--VPKEGSGNCIDYASSGNPE 178
 Db 142 VVHPLKSLGRKKNAVYVSSLVWALVVAIVADIPFYSGTGVKRNKTIICYDTTADAYLR 201
 QY 179 HNLISLCLTLGLFLIPLSYMCFYYKMWVFLKRRSQOATAPL-DKPORLVVLAIVF 237

```

; FILE REFERENCE: PC10959AGPR
; CURRENT APPLICATION NUMBER: US/10/023,775B
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: GB 0030854.4
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,590
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/296,660
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: GB 0111031.1
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-023-775B-2

Query Match 25.8%; Score 426; DB 9; Length 337;
Best Local Similarity 34.4%; Pred. No. 4.5e-30;
Matches 101; Conservative 57; Mismatches 126; Indels 10; Gaps 6;

QY 17 LNKYLSAFYAIEFICLLGNVTTFVFGYLCMKNNSSNVYLFNLSISDFAFCLTLPILI 76
Db 31 LKMHYLPVYIGIIFLVGPGNAVISTYIFKMPKSSIIIMLNACTDLYLTSLPFLI 90
QY 77 KSYAN-DKGYGVLCISNRYVLTNLTYSILFTTISMDRYLLMKYPPREHFLQKKEFA 135
Db 91 HYASGENWIFGDFMCKFIHFSEFHNLSYILFTCFIPRYCVIIHPMSCFSIHKTRCA 150
QY 136 ILISLAWALVTLEVPMLTFINSVPKEEGSNCIDYASSGNPEHNL--YSLCLTLGLFL 193
Db 151 WVACAVVMIISLVAIVPTMTLTSTNTRNSACLDLTSS--DELNIKWNLLLTATTC 208
QY 194 IPLSVMCFFYKMYVFLKRRSQQATAPLDKFORVVLAVWIFSIILFTPYHIMNLR 253
Db 209 LPLVIVILCYTIIHTLTHGLQTDSC--KOKARRLTILLAFYVGFPLPFLIRVIRIE 266
QY 254 SRLDSWPGQCT-QKAISYIYTLRPLAFINSAINPIFYFLMGDHYREMLISKFR 306
Db 267 SRLLS--ISCSIEQIHEAYIVSRPLAALTFTGNLLLYVVVSNFQAVCVSTVR 318

RESULT 10
US-10-270-144-2
; Sequence 2, Application US/10270144
; Publication No. US20030049790A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000750CON
; CURRENT APPLICATION NUMBER: US/10/270,144
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/205,196
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
; US-10-270-144-2

Query Match 25.8%; Score 426; DB 9; Length 337;
Best Local Similarity 34.4%; Pred. No. 4.5e-30;
Matches 101; Conservative 57; Mismatches 126; Indels 10; Gaps 6;

QY 17 LNKYLSAFYAIEFICLLGNVTTFVFGYLCMKNNSSNVYLFNLSISDFAFCLTLPILI 76
Db 31 LKMHYLPVYIGIIFLVGPGNAVISTYIFKMPKSSIIIMLNACTDLYLTSLPFLI 90

```

```

; FILE REFERENCE: PC10959AGPR
; CURRENT APPLICATION NUMBER: US/10/023,775B
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: GB 0030854.4
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,590
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/296,660
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: GB 0111031.1
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-023-775B-2

Query Match 25.8%; Score 426; DB 9; Length 337;
Best Local Similarity 34.4%; Pred. No. 4.5e-30;
Matches 101; Conservative 57; Mismatches 126; Indels 10; Gaps 6;

QY 17 LNKYLSAFYAIEFICLLGNVTTFVFGYLCMKNNSSNVYLFNLSISDFAFCLTLPILI 76
Db 31 LKMHYLPVYIGIIFLVGPGNAVISTYIFKMPKSSIIIMLNACTDLYLTSLPFLI 90
QY 77 KSYAN-DKGYGVLCISNRYVLTNLTYSILFTTISMDRYLLMKYPPREHFLQKKEFA 135
Db 91 HYASGENWIFGDFMCKFIHFSEFHNLSYILFTCFIPRYCVIIHPMSCFSIHKTRCA 150
QY 136 ILISLAWALVTLEVPMLTFINSVPKEEGSNCIDYASSGNPEHNL--YSLCLTLGLFL 193
Db 151 WVACAVVMIISLVAIVPTMTLTSTNTRNSACLDLTSS--DELNIKWNLLLTATTC 208
QY 194 IPLSVMCFFYKMYVFLKRRSQQATAPLDKFORVVLAVWIFSIILFTPYHIMNLR 253
Db 209 LPLVIVILCYTIIHTLTHGLQTDSC--KOKARRLTILLAFYVGFPLPFLIRVIRIE 266
QY 254 SRLDSWPGQCT-QKAISYIYTLRPLAFINSAINPIFYFLMGDHYREMLISKFR 306
Db 267 SRLLS--ISCSIEQIHEAYIVSRPLAALTFTGNLLLYVVVSNFQAVCVSTVR 318

RESULT 10
US-10-270-144-2
; Sequence 2, Application US/10270144
; Publication No. US20030049790A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000750CON
; CURRENT APPLICATION NUMBER: US/10/270,144
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/205,196
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
; US-10-270-144-2

Query Match 25.8%; Score 426; DB 9; Length 337;
Best Local Similarity 34.4%; Pred. No. 4.5e-30;
Matches 101; Conservative 57; Mismatches 126; Indels 10; Gaps 6;

QY 17 LNKYLSAFYAIEFICLLGNVTTFVFGYLCMKNNSSNVYLFNLSISDFAFCLTLPILI 76
Db 31 LKMHYLPVYIGIIFLVGPGNAVISTYIFKMPKSSIIIMLNACTDLYLTSLPFLI 90

```


Db 177 KARRLTILLLLAFYVCFPLFHILRVIRIESRLLS--ISCSTENQIHEAVIYVSRPLAALNT 234

Qy 284 AINPIFYFLMGDHYREMLISKFR 306

Db 235 FGNLLLYVVSDNFQAVCVSTVR 257

RESULT 14

US-09-789-482-4

; Sequence 4, Application US/09789482

; Patent No. US20020054875A1

; GENERAL INFORMATION:

; APPLICANT: Koch, Alisa E.

; APPLICANT: Ruth, Jeffrey H.

; APPLICANT: Rotman, James B.

; TITLE OF INVENTION: Therapeutic Methods That Target

; FILE OF INVENTION: Fractalkine or CX3CR1

; FILE REFERENCE: 3238.1000-003

; CURRENT APPLICATION NUMBER: US/09/789,482

; CURRENT FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: US 60/183,568

; PRIOR FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 355

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-789-482-4

Query Match 21.4%; Score 353; DB 10; Length 355;

Best Local Similarity 28.1%; Pred. No. 1.3e-23;

Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;

Qy 16 ILNKYILSAFYATIEFIFGLLGNVTYVFGYLFCKMKNWSSNVYLFNLSISDFAFLCTLPIL 75

Db 28 VFGTVFLSIFYVIFAGLGVNLLVVFALTNSKKPKSVTDIYLLNLALSDDLFLVATLPFW 87

Qy 76 IKSANDKGYGDVLCISNRYVLTNLTFTFISMDRYLLMKYPFREFHLOKKEFA 135

Db 88 THYLINEKGLH-NAMCKFTTAFFIGFGSIFFTVISIDRYLAIVLAANSNNRTVQHG 146

Qy 136 ILISLAWALVTLEVLPMLTFFINSVPKEGNCI-DYASSGNPEHNLIIYSLCTLLGLFLI 194

Db 147 VTISLGWAAAILVAAPQPMF----TKQENECIGDYPEVLQEIWPVLRNVETNFGFL 202

Qy 195 PLSVMCFYKMW-VFLKRSQOQATPLDKPQRLVLAIVFISILFTPYHIMRNLRIA 253

Db 203 PLLIMSYCFRIIQTFLSCNHHKAKAI-----KLILLVIVFFLFTWPNVMIFFLETL 256

Qy 254 SRLDWPQGCOTKAIKSIYTLTRPLAFNSAINPIFYFLMGDHYREMLISKFRQYFKSL 312

Db 257 KLYDFFPCDMRKDLRLALSVTETVAFSHCCNPLIYAFAGEKFRYL---YHLYGKCL 312

Search completed: April 2, 2003, 15:19:49

Job time : 23 secs

Db 177 KARRLTILLLLAFYVCFPLFHILRVIRIESRLLS--ISCSTENQIHEAVIYVSRPLAALNT 234

Qy 284 AINPIFYFLMGDHYREMLISKFR 306

Db 235 FGNLLLYVVSDNFQAVCVSTVR 257

RESULT 15

US-09-789-486-4

; Sequence 4, Application US/09789486

; Patent No. US20020055456A1

; GENERAL INFORMATION:

; APPLICANT: Koch, Alisa E.

; TITLE OF INVENTION: Therapeutic Methods That Target

; FILE OF INVENTION: Fractalkine or CX3CR1

; FILE REFERENCE: 3238.1000-004

; CURRENT APPLICATION NUMBER: US/09/789,486

; CURRENT FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: US 60/183,568

; PRIOR FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 355

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-789-486-4

Query Match 21.4%; Score 353; DB 10; Length 355;

Best Local Similarity 28.1%; Pred. No. 1.3e-23;

Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;

Qy 16 ILNKYILSAFYATIEFIFGLLGNVTYVFGYLFCKMKNWSSNVYLFNLSISDFAFLCTLPIL 75

Db 28 VFGTVFLSIFYVIFAGLGVNLLVVFALTNSKKPKSVTDIYLLNLALSDDLFLVATLPFW 87

Qy 76 IKSANDKGYGDVLCISNRYVLTNLTFTFISMDRYLLMKYPFREFHLOKKEFA 135

Db 88 THYLINEKGLH-NAMCKFTTAFFIGFGSIFFTVISIDRYLAIVLAANSNNRTVQHG 146

Qy 136 ILISLAWALVTLEVLPMLTFFINSVPKEGNCI-DYASSGNPEHNLIIYSLCTLLGLFLI 194

Db 147 VTISLGWAAAILVAAPQPMF----TKQENECIGDYPEVLQEIWPVLRNVETNFGFL 202

Qy 195 PLSVMCFYKMW-VFLKRSQOQATPLDKPQRLVLAIVFISILFTPYHIMRNLRIA 253

Db 203 PLLIMSYCFRIIQTFLSCNHHKAKAI-----KLILLVIVFFLFTWPNVMIFFLETL 256

Qy 254 SRLDWPQGCOTKAIKSIYTLTRPLAFNSAINPIFYFLMGDHYREMLISKFRQYFKSL 312

Db 257 KLYDFFPCDMRKDLRLALSVTETVAFSHCCNPLIYAFAGEKFRYL---YHLYGKCL 312

Db 177 KARRLTILLLLAFYVCFPLFHILRVIRIESRLLS--ISCSTENQIHEAVIYVSRPLAALNT 234

Qy 284 AINPIFYFLMGDHYREMLISKFR 306

Db 235 FGNLLLYVVSDNFQAVCVSTVR 257

RESULT 15

US-09-789-486-4

; Sequence 4, Application US/09789486

; Patent No. US20020055456A1

; GENERAL INFORMATION:

; APPLICANT: Koch, Alisa E.

; TITLE OF INVENTION: Therapeutic Methods That Target

; FILE OF INVENTION: Fractalkine or CX3CR1

; FILE REFERENCE: 3238.1000-004

; CURRENT APPLICATION NUMBER: US/09/789,486

; CURRENT FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: US 60/183,568

; PRIOR FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 355

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-789-486-4

Query Match 21.4%; Score 353; DB 10; Length 355;

Best Local Similarity 28.1%; Pred. No. 1.3e-23;

Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;

Qy 16 ILNKYILSAFYATIEFIFGLLGNVTYVFGYLFCKMKNWSSNVYLFNLSISDFAFLCTLPIL 75

Db 28 VFGTVFLSIFYVIFAGLGVNLLVVFALTNSKKPKSVTDIYLLNLALSDDLFLVATLPFW 87

Qy 76 IKSANDKGYGDVLCISNRYVLTNLTFTFISMDRYLLMKYPFREFHLOKKEFA 135

Db 88 THYLINEKGLH-NAMCKFTTAFFIGFGSIFFTVISIDRYLAIVLAANSNNRTVQHG 146

Qy 136 ILISLAWALVTLEVLPMLTFFINSVPKEGNCI-DYASSGNPEHNLIIYSLCTLLGLFLI 194

Db 147 VTISLGWAAAILVAAPQPMF----TKQENECIGDYPEVLQEIWPVLRNVETNFGFL 202

Qy 195 PLSVMCFYKMW-VFLKRSQOQATPLDKPQRLVLAIVFISILFTPYHIMRNLRIA 253

Db 203 PLLIMSYCFRIIQTFLSCNHHKAKAI-----KLILLVIVFFLFTWPNVMIFFLETL 256

Qy 254 SRLDWPQGCOTKAIKSIYTLTRPLAFNSAINPIFYFLMGDHYREMLISKFRQYFKSL 312

Db 257 KLYDFFPCDMRKDLRLALSVTETVAFSHCCNPLIYAFAGEKFRYL---YHLYGKCL 312

Search completed: April 2, 2003, 15:19:49

Job time : 23 secs

Db 177 KARRLTILLLLAFYVCFPLFHILRVIRIESRLLS--ISCSTENQIHEAVIYVSRPLAALNT 234

Qy 284 AINPIFYFLMGDHYREMLISKFR 306

Db 235 FGNLLLYVVSDNFQAVCVSTVR 257

RESULT 15

US-09-789-486-4

; Sequence 4, Application US/09789486

; Patent No. US20020055456A1

; GENERAL INFORMATION:

; APPLICANT: Koch, Alisa E.

; TITLE OF INVENTION: Therapeutic Methods That Target

; FILE OF INVENTION: Fractalkine or CX3CR1

; FILE REFERENCE: 3238.1000-004

; CURRENT APPLICATION NUMBER: US/09/789,486

; CURRENT FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: US 60/183,568

; PRIOR FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 355

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-789-486-4

Query Match 21.4%; Score 353; DB 10; Length 355;

Best Local Similarity 28.1%; Pred. No. 1.3e-23;

Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;

Qy 16 ILNKYILSAFYATIEFIFGLLGNVTYVFGYLFCKMKNWSSNVYLFNLSISDFAFLCTLPIL 75

Db 28 VFGTVFLSIFYVIFAGLGVNLLVVFALTNSKKPKSVTDIYLLNLALSDDLFLVATLPFW 87

Qy 76 IKSANDKGYGDVLCISNRYVLTNLTFTFISMDRYLLMKYPFREFHLOKKEFA 135

Db 88 THYLINEKGLH-NAMCKFTTAFFIGFGSIFFTVISIDRYLAIVLAANSNNRTVQHG 146

Qy 136 ILISLAWALVTLEVLPMLTFFINSVPKEGNCI-DYASSGNPEHNLIIYSLCTLLGLFLI 194

Db 147 VTISLGWAAAILVAAPQPMF----TKQENECIGDYPEVLQEIWPVLRNVETNFGFL 202

Qy 195 PLSVMCFYKMW-VFLKRSQOQATPLDKPQRLVLAIVFISILFTPYHIMRNLRIA 253

Db 203 PLLIMSYCFRIIQTFLSCNHHKAKAI-----KLILLVIVFFLFTWPNVMIFFLETL 256

Qy 254 SRLDWPQGCOTKAIKSIYTLTRPLAFNSAINPIFYFLMGDHYREMLISKFRQYFKSL 312

Db 257 KLYDFFPCDMRKDLRLALSVTETVAFSHCCNPLIYAFAGEKFRYL---YHLYGKCL 312

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 15:12:31 ; Search time 26 Seconds

(without alignments)
1172.100 Million cell updates/sec

Title: US-09-891-138A-2

Perfect score: 1650

Sequence: 1 MAQNLSCENWLATEAILNKY.....REMLSKPRQVFKSLTSFRT 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

(Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	474	28.7	373	2 JC4162	P2Y receptor - bov
2	474	28.7	373	2 JC4737	G protein-coupled
3	472.5	28.6	362	2 S33733	G protein-coupled
4	373.5	22.6	373	2 A47556	ATP receptor P2u -
5	353	21.4	355	2 JC4304	orphan G protein-c
6	352	21.3	365	2 S68679	G protein-coupled
7	336.5	20.4	360	2 A57160	chemokine (C-C) re
8	336	20.4	328	2 I55450	G protein-coupled
9	335.5	20.3	375	2 A54946	P-2U nucleotide re
10	333	20.2	308	2 I50241	G protein-coupled
11	332.5	20.2	354	2 I58186	probable G protein
12	332	20.1	328	2 JC4800	P2Y6 receptor - hu
13	328	19.9	355	2 A45177	chemokine (C-C) re
14	326.5	19.8	360	2 JC4587	chemokine (C-C) re
15	325.5	19.7	388	2 JN0605	somatostatin recep
16	318.5	19.3	359	2 A48921	interleukin-8 rece
17	317.5	19.2	355	2 A55733	interleukin-8 rece
18	316.5	19.2	384	2 A47249	G protein-coupled
19	315.5	19.1	355	2 JQ1231	brain-specific som
20	315	19.1	355	2 JC5067	interleukin-8 rece
21	314	19.0	369	2 JC2083	G protein-coupled
22	314	19.0	369	2 B41795	somatostatin recep
23	314	19.0	370	2 JC5549	somatostatin recep
24	313	19.0	344	2 T09508	hepaticellular P2Y5-
25	312.5	18.9	360	2 A53611	intron 17 purinerg
26	311	18.8	358	2 A53752	interleukin-8 rece
27	309.5	18.8	384	2 JC4629	somatostatin recep
28	309	18.7	355	2 I49339	macrophage inflam
29	308	18.7	369	2 D41795	somatostatin recep

30	307.5	18.6	369	2 A45291	somatostatin recep
31	306	18.5	356	2 I49340	MIP-1 alpha recep
32	306	18.5	362	2 A39714	G protein-coupled
33	306	18.5	391	2 A39297	somatostatin recep
34	305	18.5	346	2 S29248	somatostatin recep
35	304.5	18.5	361	2 B45680	G protein-coupled
36	303	18.4	350	2 A39445	interleukin-8 rece
37	303	18.4	391	2 A41795	somatostatin recep
38	303	18.4	391	2 C41795	somatostatin recep
39	302	18.3	366	1 JC0872	bradykinin recep
40	301.5	18.3	369	2 JC5068	G protein-coupled
41	301	18.2	41	2 G02436	chemokine (C-C) re
42	299.5	18.2	353	2 A53858	bradykinin recep
43	298	18.1	354	2 T09553	G protein-coupled
44	297	18.0	387	2 I69202	G protein-coupled
45	295.5	17.9	363	2 I57955	somatostatin recep

ALIGNMENTS

RESULT 1

JC4162

P2Y receptor - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 12-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999

C:Accession: JC4162

R:Henderson, D.J.; Elliot, D.G.; Smith, G.M.; Webb, T.E.; Dainty, I.A.

Biochem. Biophys. Res. Commun. 212, 648-656, 1995

A:Title: Cloning and characterisation of a bovine P2Y receptor.

A:Reference number: JC4162; MUID:95352058; PMID:7626079

A:Accession: JC4162

A:Molecule type: mRNA

A:Residues: 1-373 <HEND>

A:Cross-references: EMBL:X87628; NID:g1032484; PIDN:CAA60958.1; PID:g1032485

A:Experimental source: aortic endothelial cell

C:Genetics:

A:Gene: bopv2y

C:Superfamily: ATP receptor P2u

C:Keywords: glycoprotein; phosphoprotein; receptor; transmembrane protein

F:52-77/Domain: transmembrane #status predicted <TM1>

F:88-111/Domain: transmembrane #status predicted <TM2>

F:124-150/Domain: transmembrane #status predicted <TM3>

F:171-191/Domain: transmembrane #status predicted <TM4>

F:214-237/Domain: transmembrane #status predicted <TM5>

F:261-282/Domain: transmembrane #status predicted <TM6>

F:305-328/Domain: transmembrane #status predicted <TM7>

F:11,27,113,197/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:258/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

Query Match 28.7% Score 474; DB 2; Length 373;

Best Local Similarity 36.6%; pred. No. 4.2e-32;

Matches 105; Conservative 58; Mismatches 116; Indels 8; Gaps 6;

QY 20 YLSAFAYAEIFGLGNVTYVFGYFCMKNNSSNYLFNLSIDFAFCTLPILIKSY 79

Db 52 YLPAVYLIVLFLIGNSVAIMFVFMFKPWSGISVYMENALADFLYVLTIPALIFY 111

QY 80 ANDKG-TYGDVLCISNRYLHTNLYTSILFTFISMDRYLLMKYPFHEFLQKKEFAILI 138

Db 112 FNKTDWIFGDAMCKLQRFHFVNLVLSILFTICTISAHRYSGVYVPLKSLGRKKNAVYI 171

QY 139 SLAYKALVLEVLPMFLFINS-VPEKSGSNICIDYASSGNEHNLIIYSLGTLTGLFLIPLS 197

Db 172 SVLVWLVVVGISIPILYSYGIGIRKNTITCYDTSDEYLRYSFYISMCTTAMFCVPIIV 231

QY 198 VMCFEYKVMVFLKRRSQOQATALPL-DKQRLVWLVAVIFSLFTPYHIMNLRNLRASRL 256

Db 232 LILGCYGLIVRALIYKLDLNS---PLRRKSIYLVIVLVFAVSYPFFVFMKTMRLRL 288

QY 257 D-SWPGCT-QKAIKSYTYTTRPIAFNLNSAINPIFYFLMGDHYREML 301

Db 289 DFQTPENCAFNDRYATYQVTRGLASLNSCVDPILYFLAGDTFRRL 335

RESULT 2

JC4737
G protein-coupled receptor P2Y1 - human
N:Alternate names: P2Y1 purinergic receptor; P2Y1 purinoreceptor
C:Species: Homo sapiens (man)
C:Date: 10-May-1996 #sequence_revision 16-Aug-1996 #text_change 17-Nov-2000
C:Accession: JC4737; J04615; S54253
R:Janssens, R.; Communi, D.; Pirotton, S.; Samson, M.; Parmentier, M.; Boeynaems, J.M.
Biochem. Biophys. Res. Commun. 221, 588-593, 1996
A:Title: Cloning and tissue distribution of the human P2Y1 receptor.
A:Reference number: JC4737; MUID:96205320; PMID:8630005
A:Accession: JC4737

A:Molecule type: DNA
A:Residues: 1-373 <JAN>
A:Cross-references: GB:S81950; NID:gl1839438; PIDN:ABA47091.1; PID:gl1839439
R:Ayyanathan, K.; Webbs, T.E.; Sandhu, A.K.; Athwal, R.S.; Barnard, E.A.; Kunapuli, S.P.
Biochem. Biophys. Res. Commun. 218, 783-788, 1996
A:Title: Cloning and chromosomal localization of the human P2Y1 purinoreceptor.
A:Reference number: JC4615; MUID:96158962; PMID:8579591
A:Accession: JC4615

A:Molecule type: mRNA
A:Residues: 1-373 <AYY>
A:Cross-references: GB:U04209; NID:gl147730; PIDN:AAA97872.1; PID:gl147731
A:Experimental source: erythro leukemia cells
R:Leon, C.; Vial, C.; Cazenave, J.; Gachet, C.
submitted to the EMBL Data Library, May 1995
A:Description: Cloning of a human putative P2Y receptor.
A:Reference number: S54253
A:Accession: S54253

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-137, 139-373 <LEO>
A:Cross-references: EMBL:249205; NID:g798835; PIDN:CAA9066.1; PID:g798836
C:Comment: This receptor belongs to a family of G protein-coupled receptors. It responds
C:Genetics:
A:Gene: P2Y1; GDB:P2Y1
A:Cross-references: GDB:677125; OMIM:601167

A:Map position: 3pter-3qter
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
F:52-77/Domain: transmembrane #status predicted <TM1>
F:88-111/Domain: transmembrane #status predicted <TM2>
F:124-152/Domain: transmembrane #status predicted <TM3>
F:171-191/Domain: transmembrane #status predicted <TM4>
F:214-237/Domain: transmembrane #status predicted <TM5>
F:261-282/Domain: transmembrane #status predicted <TM6>
F:305-328/Domain: transmembrane #status predicted <TM7>

F:11,27,113,197/Binding site: carboxylate (Asn) (covalent) #status predicted
F:258,336/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted
F:330,339/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:343/Binding site: phosphate (Ser) (covalent) (by protein kinase C and calmodulin-depen

Query Match 28.7%; Score 474; DB 2; Length 373;
Best Local Similarity 37.3%; Pred. No. 4.2e-32;
Matches 107; Conservative 55; Mismatches 117; Indels 8; Gaps 6;

QY 20 YLSAFYAIEFIFGLGNVTVFVGLFCMKNNSSNVFLNLSISDFALCTLPILIKSY 79
DB 52 YLPNAVILVFIIGFLGNSVAIMWFVFMKPSGIVSYFENALADFLVLTLPALIFVY 111
QY 80 ANDKG-TYGDVLCISNRYLHTNLYTSILFTFISMDRYLLMKYPPREHFLQKFAILI 138
DB 112 FNKTDWIFGDAMCKLQRFIFHVNLYGSLFTCTSAHRSYGVVYPLKSLGRLLKNAICI 171
QY 139 SLAWVALVTEVLPMLTFINS-VPKESGNCIDYASSGNFHNLYSLCTLTLLGLFPLIS 197
DB 172 SVLVMLIVVAISPIFYSGTVGRVKNKNTITCYDTSDYLRISFYISMCVTVMFCVPLV 231
QY 198 VMCFYFKMVFVFLKRRSQOQATALPL-DKQRLVVLAVVIFSLFTPYHMRNLRTASRL 256
DB 232 LILGCYGLIVRALIYKDLNDS---PLRRKSYLIVLTVFAVSYIPFHVMTNLRRL 288

QY 257 D-SWQOQCT-QKAIKSIYTLTRPLAFNSAINPIFYFLMGDHYREML 301
DB 289 DFQTPAMCAFNDRYATYQVTRGLASLNSCVDPIYFLAGDTFRRL 335

RESULT 3

S33733
G protein-coupled receptor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999
C:Accession: S33733
R:Webb, T.E.; Simon, J.; Krishek, B.J.; Bateson, A.N.; Smart, T.G.; King, B.F.; Burns
FEBS Lett. 324, 219-225, 1993
A:Title: Cloning and functional expression of a brain G-protein-coupled ATP receptor.
A:Reference number: S33733; MUID:93285340; PMID:8508924
A:Accession: S33733

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-362 <WEB>
A:Cross-references: EMBL:X73268; NID:g395084; PIDN:CAA51716.1; PID:g395085
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 28.6%; Score 472.5; DB 2; Length 362;
Best Local Similarity 35.9%; Pred. No. 5.4e-32;
Matches 110; Conservative 58; Mismatches 121; Indels 17; Gaps 8;

QY 10 WLA----TEAILNK-----YLSAFYAIEFIFGLGNVTVFVGLFCMKNNSSNVYLEN 60
DB 22 WAGNATTKCSLTKTGFQFYVLTPTVYILFTFGLGNSVAIMWFVFMKPSGIVSYFEN 81
QY 61 LSISDFALCTLPILIKSYANDKG-TYGDVLCISNRYLHTNLYTSILFTFISMDRYLL 119
DB 82 LALADFLVLTLPALIFYFNKTDWIFGDVCKLQRFIFHVNLYGSLFTCTLSVHRYTG 141
QY 120 MKYPERHFLOKKEFAILISLAVWALVTLVPLMLTFINS-VPKESGNCIDYASSGNPE 178
DB 142 VVHLKSLGRLKKNVAVVSSVSLVAVVAVIAPILFYSGTVGRVKNKNTITCYDTTAD 201
QY 179 HNLYSLCTLTLLGLPLSVCMFFYKMYVFLKRRSQOQATALPL-DKQRLVVLAVVIF 237
DB 202 SYFVSMCTTTFMFCPIFIVLIGCYGLIVKALYKDLNDS---PLRRKSYLIVLTVTF 258
QY 238 SILTPYHMRNLRTASRLD-SWQOQCT-QKAIKSIYTLTRPLAFNSAINPIFYFLMGD 295
DB 259 AVSYLPFHVNKTLNLRKLDPTQFMCAFNDKVVATYQVTRGLASLNSCVDPIYFLAGD 318
QY 296 HYREML 301
DB 319 TFRRL 324

RESULT 4

A47556
ATP receptor P2u - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C:Accession: A47556
R:Lustig, K.D.; Shiau, A.K.; Brake, A.J.; Julius, D.
Proc. Natl. Acad. Sci. U.S.A. 90, 5113-5117, 1993
A:Title: Expression cloning of an ATP receptor from mouse neuroblastoma cells.
A:Reference number: A47556; MUID:93281707; PMID:7685114
A:Accession: A47556
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-373 <LUS>
A:Cross-references: GB:L14751; NID:g309457; PIDN:AAA39871.1; PID:g309458
C:Superfamily: ATP receptor P2u
C:Keywords: transmembrane protein

Query Match 22.6%; Score 373.5; DB 2; Length 373;
Best Local Similarity 33.5%; Pred. No. 1e-23;

Matches 94; Conservative 50; Mismatches 130; Indels 7; Gaps 5;

QY 19 KY-YLSAPYATIEFGILGNVTVVFGYLCMKWNSSNYLFLNLSISDFAFICTLPILIK 77
 Db 32 KYVLLPSYGVVGVGLGNVVALYIFLCRLKTNASTTMYFHLAVSDSLYAAASPLLYV 91
 QY 78 KYA-NDKGTGVDVLCISNRYVLTNLTSTLFTFISMDRYLLMKYPFREFHFLQKEFAI 136
 Db 92 YYARGDHPFSTVLCKLVRELVFNLYCSILFTLCISVHRCLGVRLPLSLRWGRARYAR 151
 QY 137 LISLAWALVTLVPLMLTINSVPKEGNCIDYASSGNPHENLIYSCLLGLFLPL 196
 Db 152 RVAAVWVLVLAQAPVLYIEVTVSVRGTRITCHDTSARELFHFVAYSSVMLGLLFAVPF 211
 QY 197 SVMCFYKVMVFLKRRSQOATALPLDKPQ--RLVVLAVLFSILFTPYHIMRNLRAS 254
 Db 212 SVILVCVLMARLLRPAYGTTGGLPRARKSVRTIALVLAFCFLDFPHVTRITLYYSF 271
 QY 255 RLDSWPGCC--TQKAISKIYTLRPLAFINSAINPIFYFLMG 294
 Db 272 R--SLDLSCHTLAINMAYKITRPLASANSCLDPVLYFLAG 310

RESULT 5

JC4304
 orphan G protein-coupled receptor - human
 N:Alternate names: V28 protein
 C:Species: Homo sapiens (man)
 C:Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 19-May-2000
 C:Accession: JC4304
 R:Report, C.J.; Schweickart, V.L.; Eddy Jr., R.L.; Shows, T.B.; Gray, P.W.
 Gene 163, 295-299, 1995
 A:Title: The orphan G-protein-coupled receptor-encoding gene V28 is closely related to 9
 A:Reference number: JC4304; MUID:96011651; PMID:7590284
 A:Accession: JC4304
 A:Molecule type: mRNA
 A:Residues: 1-355 <GB>
 A:Cross-references: GB:U20350; NID:9665580; PIDN:AA91783.1; PID:9665581
 A:Experimental source: peripheral blood mononuclear cell
 C:Comment: This protein is a cell-surface receptor which recognizes extracellular signal
 C:Genetics:
 A:Gene: v28
 A:Map position: 3pter-p21
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; lymphokine; transmembrane protein
 F:35-37/Domain: transmembrane #status predicted <TM1>
 F:66-88/Domain: transmembrane #status predicted <TM2>
 F:104-125/Domain: transmembrane #status predicted <TM3>
 F:146-165/Domain: transmembrane #status predicted <TM4>
 F:197-217/Domain: transmembrane #status predicted <TM5>
 F:230-254/Domain: transmembrane #status predicted <TM6>
 F:275-296/Domain: transmembrane #status predicted <TM7>

Query Match 21.4%; Score 353; DB 2; Length 355;
 Best Local Similarity 28.1%; Pred. No. 5e-22;
 Matches 84; Conservative 66; Mismatches 133; Indels 16; Gaps 6;

QY 16 ILNKYLSAFYATIEFGILGNVTVVFGYLCMKWNSSNYLFLNLSISDFAFICTLPIL 75
 Db 28 VFGTFLSIFSYVFAIGLVGLNLLVFAITNSKKPKSVTDIYLLNLALSDDLLEFVATLPWF 87
 QY 76 IKSYANDKGTGVDVLCISNRYVLTNLTSTLFTFISMDRYLLMKYPFREFHFLQKEFA 135
 Db 88 THYLINERGLH-NAMCKETTAFFGIFGSGIFFITVISIDRYLAIVLAANSMMNRVQHG 146
 QY 136 ILISLAWALVTLVPLMLTINSVPKEGNCIT-DYASSGNPHENLIYSCLLGLFLI 194
 Db 147 VTISLVWAAAILVAAPQFMF----TKOKENECLGDYPEVLQETWPLRVLRNVEVNFGLFL 202
 QY 195 PLSVCMCFYKVMV-VFLKRRSQOQATALPLDKKQRLVLAWVLSILFTPYHIMRNLRAS 253
 Db 203 PLLIMSYCFRIIOTLPSCKNHKKAKAI-----KLILLVIVFFLFTPYPNVWIFLETL 256

QY 254 SRLDSWPGCCTQKAISKIYTLRPLAFINSAINPIFYFLMGDHYRMLISKRFQYFKSL 312
 Db 257 KLYDFPSCDMRDLRLALSVTETVAFSHCCCLNPLIYAAGEKFRYL---YHLYGKCL 312
 RESULT 6
 S68679
 G protein-coupled receptor - human
 C:Species: Homo sapiens (man)
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
 C:Accession: S68679
 R:Stam, N.J.; Klomp, J.; van de Heuvel, M.; Olijve, W.
 FEBS Lett. 384, 260-264, 1996
 A:Title: Molecular cloning and characterization of a novel orphan receptor (P(2P)) ex
 A:Reference number: S68679; MUID:96197801; PMID:8617367
 A:Accession: S68679
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-365 <STA>
 A:Cross-references: EMBL:X96597; NID:gl296631; PIDN:CAA65415.1; PID:gl296632
 C:Superfamily: ATP receptor P2u
 C:Keywords: G protein-coupled receptor

Query Match 21.3%; Score 352; DB 2; Length 365;
 Best Local Similarity 31.1%; Pred. No. 6.2e-22;
 Matches 90; Conservative 59; Mismatches 122; Indels 18; Gaps 6;
 QY 22 LSAYFAIEFIFGLGNVTVVFGYLCMKWNSSNYLFLNLSISDFAFICTLPILIKSY-A 80
 Db 38 LPYSYAVVFGVGLGNLNAPTLWLFIFRLRPWDATATYMFHLASDPLVYLSLPTLIYYAA 97
 QY 81 NDKGTGVDVLCISNRYVLTNLTSTLFTFISMDRYLLMKYPFREFHFLQKEFAILISL 140
 Db 98 HHWPFGTEICKFVRFLFYNNLYCSVLFITCISVHRVILGICHPRLALRWGRPRLAGLCL 157
 QY 141 AVNALVTLEVLPMILFINSVPKEGNCIDYASSGNPHENLIYSCLLPLGLFLPLSVMC 200
 Db 158 AVNLVAVAGCLVPLNFVITTSNKGTTVLCHDTRPEFDHYVHFSSAVMGLLFGVPCVLTL 217
 QY 201 FFYKVMVFLKRR-----SQQATALPLDKQRLVLAWVLSILFTPYHIMRNLRAS 253
 Db 218 VCYGLMA----RRLYQPLFGSAQSSRL---RSLRTIAVLTVEAVCFVPHITRTIYVL 270
 QY 254 SRLDSWPGCC--TQKAISKIYTLRPLAFINSAINPIFYFLMGDHYRML 301
 Db 271 ARL--LEADCRVLNIVNVVYKVTREPLASANSCLDPVLYLLTGDKYRROL 317

RESULT 7

A57160
 chemokine (C-C) receptor 4 - human
 N:Alternate names: C-C CKR-4
 C:Species: Homo sapiens (man)
 C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
 C:Accession: A57160
 R:Power, C.A.; Meyer, A.; Nemeth, K.; Bacon, K.B.; Hoogewerf, A.J.; Proudfoot, A.E.I.
 J. Biol. Chem. 270, 19495-19500, 1995
 A:Title: Molecular cloning and functional expression of a novel CC chemokine receptor
 A:Reference number: A57160; MUID:95370289; PMID:7642634
 A:Accession: A57160
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-360 <POW>
 A:Cross-references: GDB:X85740; NID:g1370103; PIDN:CAA59743.1; PID:g971452
 A:Note: source clone K5-5
 C:Genetics:
 A:Gene: GDB:CMKBR4
 A:Cross-references: GDB:677463
 A:Map position: 3p21-3p21
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane p
 F:40-65/Domain: transmembrane #status predicted <TM1>

F:76-97/Domain: transmembrane #status predicted <TM2>
F:112-133/Domain: transmembrane #status predicted <TM3>
F:151-175/Domain: transmembrane #status predicted <TM4>
F:208-226/Domain: transmembrane #status predicted <TM5>
F:243-264/Domain: transmembrane #status predicted <TM6>
F:291-308/Domain: transmembrane #status predicted <TM7>
F:299-276,110-187/disulfide bonds: #status predicted
F:72,350/Binding site: phosphate (Ser) (covalent) (by protein kinase II) #status predicted
F:145/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F:183,194/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:321/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 20.4%; Score 336.5; DB 2; Length 360;
Best Local Similarity 29.4%; Pred. No. 1.2e-20;
Matches 93; Conservative 60; Mismatches 132; Indels 31; Gaps 9;

QY 14 EATLNKYL-----SAF-----YAEIFGGLGNVTGVFGYLFCKMKNW 52
DB 13 ESIYNYLYESTPKCTKEGKAGLELPPLYSLVFVGLGNSVVVL-VLFYKRLR 71
QY 53 S-SNVYLFNLSISDFALCTLPILKSYANDKGTGYDVLICISNRVYLTNLTSLFTLF 111
DB 72 SMTDVLNLAISDLFLVESLPW-CYAAADQWVFGGLCKMISWMLVGFYSGIFFVWL 130
QY- 112 ISMDRYLLMKYPPEHFLQKKEFAILLISLAVNALVTLVPLMLTFINSVPKEGNCIDY 171
DB 131 MSIDRYLAIVHAFVSRLARTLTGYVITSLATSVAVFASLPGLFSTCTERNHYCKTK 190
QY 172 ASSGNPEHNLISGLTLGLFLGVLVSVMGFFYKVMVFLKRRSQOQATALPDKPQRLV 231
DB 191 YLSNSTTWKVLSSLEINILGLVPLGIM-LFCYSMII-----RTLQHCNKKNAVKMIF 245
QY 232 LAVVIFSLFTPYHIMRNLIARSLQSWPGQCT-QRAIKSIYTLTFLPLAFNSAINPIFY 290
DB 246 AVVFLGFETPYNIVLFTLLEVL-QDCTFERLYDAIQTATLAFVHCCLNPIIY 304
QY 291 FLMGDHVREMLISKER 306
DB 305 FFLGKERFKYILQLFK 320

RESULT 8
I55450
G protein-coupled P2 receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-May-2000
C:Accession: I55450
R:Chang, K.; Hanaka, K.; Kumada, M.; Takuwa, Y.
J. Biol. Chem. 270, 26152-26158, 1995
A:Title: Molecular cloning and functional analysis of a novel P2 nucleotide receptor.
A:Reference number: I55450; MUID:96064682; PMID:7592819
A:Accession: I55450
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <RES>
A:Cross-references: GB:D63665; NID:g1066007; PIDN:BAA09816.1; PID:g1066008
A:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor

Query Match 20.4%; Score 336.5; DB 2; Length 328;
Best Local Similarity 29.9%; Pred. No. 1.2e-20;
Matches 85; Conservative 52; Mismatches 137; Indels 10; Gaps 4;

QY 22 LSFAFAIEFICLLGNVTGVFGYLFCKMKNWNSVYLFNLSISDFALCTLPILIKSYA- 80
DB 29 LPPYVSVLVGLPLNVCVIAQICASRTLTSAVYVILNALADLLIYACSLPLIYNAR 88
QY 81 NDKGYIGDVLICISNRVYLTNLTSLFTLPIFISMDRYLLMKYPPE-HFLQKKEFAILLIS 139
DB 89 GDHWFGDLACRLVRFYANLHGSILFTLCISFORVILGICHLAPLAPWHKRGRRAAWVC 148
QY 140 LAVNALVTLVPLMLTFINSVPKEGNCIDYASSGNPEHNLISGLTLGLFLPLISVM 199

112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

149 GVWLVVTAQCLTAVFAATGIQNRNVCYDLSPPLSTRYLPGMALVIGFLPLFTAL 208
200 CFYYKMWFLKRRSQOQATLPL-----DKPQRLVVLAVVIFSLFTPYHIMRNLIAS 254
209 LACYCRMA---RRLCKRODGPAGVAOERSKAAARMVAVVAEIVISFLPHITKTAYLAV 265
255 RLDSWPQCGTQKAIKSIYTLTRELAFNLNSAINPIFYFLMGDHYR 298
266 RSPFGVSCPVEITFAAAYKGRFPASANSVLDPILEVFVFTQOKPR 309
RESULT 9
A54946
P-2U nucleotide receptor - human
C:Species: Homo sapiens (man)
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 17-Mar-1999
C:Accession: A54946
R:Par, C.E.; Sullivan, D.M.; Paradiso, A.M.; Lazarowski, E.R.; Burch, L.H.; Olsen, J.
Proc. Natl. Acad. Sci. U.S.A. 91, 3275-3279, 1994
A:Title: Cloning and expression of a human P-2U nucleotide receptor, a target for cys
A:Reference number: A54946; MUID:94211846; PMID:8159738
A:Accession: A54946
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-375 <PAR>
A:Cross-references: GB:U07225
A:Note: parts of this sequence were confirmed by protein sequencing
C:Genetics:
A:Gene: GDB:P2RY2; HP20; P2U
A:Cross-references: GDB:362713; OMIM:600041
A:Map position: l1q13.5-l1q14.1
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 20.3%; Score 335.5; DB 2; Length 375;
Best Local Similarity 33.1%; Pred. No. 1.5e-20;
Matches 93; Conservative 52; Mismatches 127; Indels 9; Gaps 7;

QY 19 KY-YLSAFYAIEFTFGLGNVTGVFGYLFCKMKNWNSVYLFNLSISDFALCTLPILIK 77
DB 32 KYVLLPVSVGVVGLCLNAVGLYIFLCRLKTNASTYMFHLAVSDALYAAASLPLVY 91
QY 78 SYA-NDKGTGYDVLICISNRVYLTNLTSLFTLPIFISMDRYLLMKYPPEHFLQKKEFAI 136
DB 92 YIARGDHPFSTVLCIKLVRFYTNLYCISILFCISVHRCGLVLRPLSLRWGRARYAR 151
QY 137 LISLAWALVTLVLEPLMTFINSVPKEGNCIDYASSGNPEHNLISGLTLGLFLIPL 196
DB 152 RVAGAVVVLVACQAPVLYFVTTSSARGP-LTCHDTSAPELFSRFVAYSSVMGLLFAVFF 210
QY 197 SVMCFYKMWVFLKRRSQOQATLPLQKQRLVVLAVV--IFSLFTPYHIMRNLIAS 254
DB 211 AVILVCYVLMARRLLKPAYGTSGGLPRAKRSKSVRTIAVVLAVFALCFLPHFVHTRLYYSF 270
QY 255 RLDSWPQCGTQKAIKSIYTLTRELAFNLNSAINPIFYFLMG 294
DB 271 R-SLDLSCHTLNAINMAYKVTR-LASANSCLDPLVLYFLAG 308

RESULT 10
I50241
G protein-coupled receptor 6H1 - chicken
N:Alternate names: purinoceptor 6H1
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Jun-2000
C:Accession: I50241; JC4618
R:Kaplan, M.H.; Smith, D.I.; Sundick, R.S.
J. Immunol. 151, 628-636, 1993
A:Title: Identification of a G protein coupled receptor induced in activated T cells.
A:Reference number: I50241; MUID:93329058; PMID:8393036
A:Accession: I50241
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

Query Match	20.2%;	Score 332.5;	DB 2;	Length 354;
Best Local Similarity	27.9%;	Pred. No. 2.6e-20;		

C: Superfamily: ATP receptor P2u
F: Keywords: glycoprotein; placenta; receptor; transmembrane protein
F:26-52/Domain: transmembrane #status predicted <TM1>
F:63-86/Domain: transmembrane #status predicted <TM2>
F:104-122/Domain: transmembrane #status predicted <TM3>
F:143-167/Domain: transmembrane #status predicted <TM4>
F:193-216/Domain: transmembrane #status predicted <TM5>
F:241-264/Domain: transmembrane #status predicted <TM6>
F:283-305/Domain: transmembrane #status predicted <TM7>
F:57-173/Binding site: carbohydrate (Asn) (covalent) status predicted

[illegible]

Db 149 VAWLAVTTQCLPTAFAATGIORNTVCYDLSPPALATHYMPYGMALTVIGFLLPFAAL 208

QY 200 CFYYKVVVFLKRRSQOQATAPL-----DKPQLVWLVAVVIFSLFTPYHIMRNLRAS 254

Db 209 LACYCLLAC---RLCRQDGPAPVQERGAARVAVVAAFAISLFPFHITKTAYLAV 265

QY 255 RLDSWPQGTQKRAIKSIYTLTRPLAFNSAINPI-FYFLMGDHYREMLISKFRQ 307

Db 266 RSTPGVPCTVLEAFAAAYKGRPFASANSVLPILFYFTQ-----KKFRR 310

RESULT 13

A45177

chemokine (C-C) receptor 1 - human

N:Alternate names: C-C CKR-1; macrophage inflammatory protein-1-alpha receptor

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999

C:Accession: A45177; 155671

R:Neote, K.; DiGregorio, D.; Mak, J.Y.; Horuk, R.; Schall, T.J.

Cell 72, 415-425, 1993

A:Title: Molecular cloning, functional expression, and signaling characteristics of a C-

A:Reference number: A45177; MUID:93161416; PMID:7679328

A:Accession: A45177

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-355 <NEO>

A:Cross-references: GB:L10918; NID:9292416; PIDN:AAA36543.1; PID:9292417

A:Experimental source: HL60 cells

A:Note: sequence extracted from NCBI backbone (NCBIP:124876)

R:Gao, J.

J. Exp. Med. 177, 1421-1427, 1993

A:Title: Structure and functional expression of the human macrophage inflammatory 1 alpha

A:Reference number: I55671; MUID:93240122; PMID:7683036

A:Accession: I55671

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-355 <RES>

A:Cross-references: GB:L10918; NID:9292416; PIDN:AAA36543.1; PID:9292417

C:Genetics:

A:Gene: GDB:CMKR1; CMKR-1

A:Cross-references: GDB:138446; OMIM:601159

A:Map position: 3p21-3p21

C:Superfamily: vertebrate rhodopsin

C:Keywords: disulfide bond; G protein-coupled receptor; glycoprotein; phosphoprotein; tr

F:36-60/Domain: transmembrane #status predicted <TM1>

F:71-91/Domain: transmembrane #status predicted <TM2>

F:108-129/Domain: transmembrane #status predicted <TM3>

F:147-171/Domain: transmembrane #status predicted <TM4>

F:205-223/Domain: transmembrane #status predicted <TM5>

F:240-264/Domain: transmembrane #status predicted <TM6>

F:288-305/Domain: transmembrane #status predicted <TM7>

F:5/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:24-273,106-183/Disulfide bonds: #status predicted

F:345/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted

Query Match 19.9%; Score 328; DB 2; Length 355;

Best Local Similarity 29.0%; Pred. No. 6.1e-20;

Matches 85; Conservative 63; Mismatches 127; Indels 18; Gaps 8;

QY 22 LSAFYAIEFIFGLGNVTVPFGYLFCKMKNWNSNVFLNLSISDFACFLCTLPILIKSYAN 81

Db 37 LPPLYSLVFVIGLVGNILVLVVOYRKLNKNTSIYLLNLAISDLLEFLTPFWIDYKX 96

QY 82 DKGTGYDGLC--ISNRYVLHNLXTSLTLFTFISMDLMLKMPFFRHFQKKEFAILIS 139

Db 97 DDWFGDAMCKILSGFY--YGLGYSEIFFIILLTIDRYLAIVHAVALRARTVTFGITS 154

QY 140 LAVWALVTLEVLPMLTFTNSVPKKEGSCN-IDYASSGNPEHNLISLCLTLGLPLPSV 198

Db 155 IIVWALAILASMPGLYFFSKTQWEFTHHFCSHFPESHREWKLFQALKLFGVLVLELV 214

QY 199 MCFEYKVVVFLKRRSQOQATAPLDPKQRLVWLVAVVIFSLFTPYHIMRNLRASRLD- 257

Db 215 MIICYGTGIKILLRRPNKKSS-----KAVRLIEFVIMIIFFLFTWPNYL--TILISVQDF 267

QY 258 SWPGCGTQ-KAIKSIYTLTRPLAFNSAINPIFYFLMGDHYREMLISKPRQYF 309

Db 268 LFTHECEQSRHLDIAVQVTEVIATHTCCVNPVIYAFVGERFRKYL-----RQLF 316

RESULT 14

JC4587

chemokine (C-C) receptor 4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 08-Mar-1996 #sequence_revision 19-Apr-1996 #text_change 20-Jun-2000

C:Accession: JC4587

R:Hoogewerf, A.J.; Black, D.; Proudfoot, A.E.I.; Wells, T.N.C.; Power, C.A.

Biochem. Biophys. Res. Commun. 218, 337-343, 1996

A:Title: Molecular cloning of murine CC CKR-4 and high affinity binding of chemokines

A:Reference number: JC4587; MUID:96136324; PMID:8573157

A:Accession: JC4587

A:Molecule type: mRNA

A:Residues: 1-360 <HOO>

A:Cross-references: EMBL:X90862; NID:g1167851; PIDN:CAA62372.1; PID:g1167852

A:Experimental source: thymus

C:Genetics:

A:Gene: cc ckr-4

C:Superfamily: vertebrate rhodopsin

C:Keywords: glycoprotein; phosphoprotein; receptor; thymus

F:2,183,194/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:72,202,350/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status p

F:145/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predi

F:321/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predi

Query Match 19.8%; Score 326.5; DB 2; Length 360;

Best Local Similarity 28.4%; Pred. No. 8.2e-20;

Matches 93; Conservative 57; Mismatches 136; Indels 41; Gaps 10;

QY 14 EAILNKYY-----LSAFYAIEFIFGLGNVTVPFGYLFCKMKNW 52

Db 13 .EVVYNSYFYFESNPKPCKEGIKAFGEVFLPPLYSLSVLLGLFGNSVVVL-VLEFKYRKLK 71

QY 53 S-SNYLFLNLSISDFACFLCTLPILIKSYANDKGYGVLCISNRYVLHNLXTSLFTTF 111

Db 72 SMTDVLNLAISDLLEFLVLSLPFW-GYAAADQWVEGLCKIVSNWYLVGYSGIFFML 130

QY 112 ISMDRYLLMKYFFRHFQKKEFAILISLWALVTLEVLPMLTFTNSVPKKEGSCNIDY 171

Db 131 MSIDRYLAIVHAVALRARTVTFGITSLSLTSWVAVFASLPGLLSTCYTEHNHTYCKTQ 190

QY 172 ASSGNPEHNLISLCLTLGLPLPSVMCFYKVMVFLKRRSQOQATAPLDPKQRLVW 231

Db 191 YSVNSTTKVLSLSLEINVLGLLPLGIM-LFWYSMII-----RTLQCKNEKKNRAVRMIF 245

QY 232 LAVVIFSLFTPYHIMRNLRASRLDSNPQGTCT-OKAIKSIYTLTRPLAFNSAINPIFY 290

Db 246 GVVLFGFWTPYVNVVFLTLEVLVLEVL-ODCTLERLYDAIQATETLGFTHCLNPVIY 304

QY 291 FLMGDHYREMLISKPRQYFKSLTSFT 317

Db 305 FFLGE-----KFKYITQL--FRT 321

RESULT 15

JN0605

somatostatin receptor 4 - human

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000

C:Accession: JN0605; JN0762; A47457

R:Xu, Y.; Song, J.; Bruno, J.F.; Berelowitz, M.

Biochem. Biophys. Res. Commun. 193, 648-652, 1993

A:Title: Molecular cloning and sequencing of a human somatostatin receptor, hSTR4.

A:Reference number: JN0605; MUID:93290656; PMID:8512564

A:Accession: JN0605

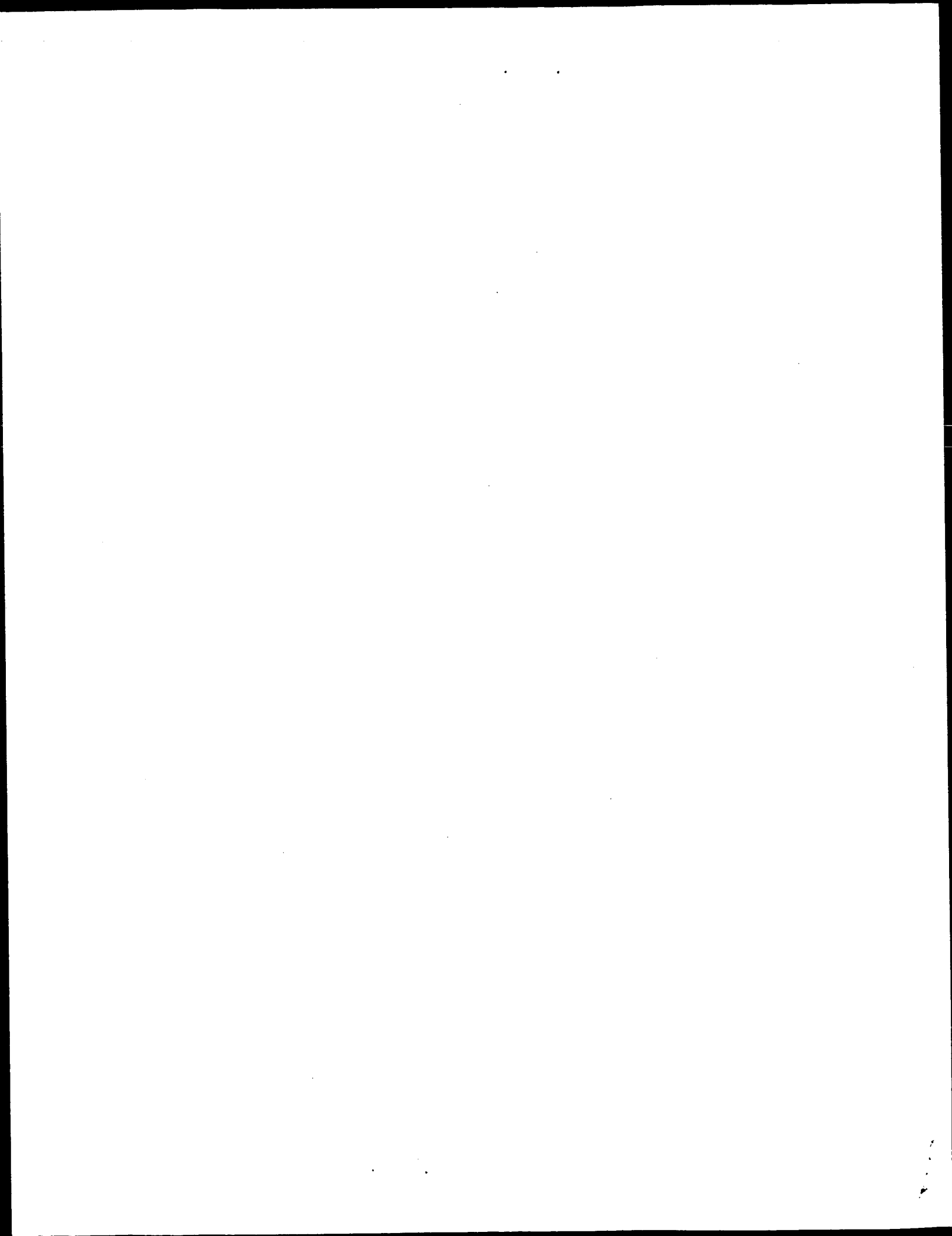
A:Molecule type: DNA

A:Residues: 1-388 <XUY>

A:Cross-references: GB:U14856; NID:9292499; PIDN:AAA36623.1; PID:g292500
R:Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; I
Biochem. Biophys. Res. Commun. 195, 844-852, 1993
A:Title: Cloning, functional expression and pharmacological characterization of a fourth
A:Reference number: JN0762; MUID:93384611; PMID:8373420
A:Accession: JN0762
A:Molecule type: DNA
A:Residues: 1-388 <YAM>
A:Cross-references: GB:D16826; NID:9693907; PIDN:BAA04106.1; PID:g693908
R:Rohrer, L.; Raulf, F.; Bruns, C.; Buettner, R.; Hofstaedter, F.; Schule, R.
Proc. Natl. Acad. Sci. U.S.A. 90, 4196-4200, 1993
A:Title: Cloning and characterization of a fourth human somatostatin receptor.
A:Reference number: A47457; MUID:93248256; PMID:8483934
A:Accession: A47457
A:Molecule type: DNA
A:Residues: 1-82, 'T', 84-364, 'K', 366-388 <ROH>
A:Cross-references: GB:L07833; NID:g307429; PIDN:AAA60565.1; PID:g307430
A:Note: sequence extracted from NCBI backbone (NCBI:130856, NCBI:130856)
C:Comment: This protein mediates the diverse actions of the tetradecapeptide somatostatin.
C:Genetics:
A:Gene: GDB:SSTR4
A:Cross-references: GDB:202662; OMIM:182454
A:Map position: 20p11.2-20p11.2
A:Introns: #status absent
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; pho
F:47-73/Domain: transmembrane #status predicted <TM1>
F:84-109/Domain: transmembrane #status predicted <TM2>
F:121-142/Domain: transmembrane #status predicted <TM3>
F:162-184/Domain: transmembrane #status predicted <TM4>
F:208-238/Domain: transmembrane #status predicted <TM5>
F:257-284/Domain: transmembrane #status predicted <TM6>
F:291-314/Domain: transmembrane #status predicted <TM7>
F:24/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:119-198/Disulfide bonds: #status predicted
F:161,253/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status pr
F:321/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match	19.7%;	Score 325.5;	DB 2;	Length 388;
.Best Local Similarity	28.4%;	Pred. No. 1.1e-19;		
Matches	82;	Conservative	66;	Mismatches 110; Indels 31; Gaps 5;
QY	22	LSAFYAIEFFIGLLGNVTVVFGYLFCKMKNSSNVYLFNISDFAELCTLPILIKSYAN	81	
DQ	50	IQCIYALVCLVUGNALVIFVILRYAKMKATATNIYLLNLADELFMLSPVFVASSAAL	109	
QY	82	DKGTYGDVLCISNRYLHTNLYTSILFTFISMDRYLLMKYPRREHFLQKKEAIIISLA	141	
Db	110	RHWPFGSVLCRAVLSVDGLNMFTSVCLTVLSVDRIYAVVHPHRAATYRSPSVAKLINLG	169	
QY	142	VWALVTLEVLPMLTFINSPKPEGSCNICDYASSGNPEHNLIIYSLCTLTLLGLFIPLSVMCF	201	
Db	170	VWLASLLVTLPFAIFADTRPARGQAVACNLQWPHPAWSAVFVVYFTLLGLFLPVLAIGL	229	
QY	202	FYKMW-----VPL-----KRSQOQATPLDQFQRLVVLAWTFSILFTPYHIMRNL	250	
Db	230	CYLLIYGMKRAVALRAGWOQRREKKIT-----RLVLMVVVYFVLCWMPFYVVQLL	281	
QY	251	R-IASRLDSWPQCCTQKAITSYTLTRPLAFNSAINPIFYFLMGDHYR	298	
Db	282	NLVVTSLSA-----TVNHVSLLIUSYANCSNPILYGLFSNFR	319	

Search completed: April 2, 2003, 15:18:04
Job time : 26 secs



GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 14:32:31 ; Search time 16 seconds
(without alignments)
821.750 Million cell updates/sec

Title: US-09-891-138a-2

Perfect score: 1650

Sequence: 1 MAQNLSNENLATEAILNK.....REMLSKPROYFKSLTSFRT 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	479	29.0	373	1	P2YR_RAT
2	475	28.8	373	1	P2YR_MOUSE
3	474	28.7	373	1	P2YR_BOVIN
4	474	28.7	373	1	P2YR_HUMAN
5	472.5	28.6	362	1	P2YR_CHICK
6	472.5	28.6	362	1	P2YR_MELEAGRIS
7	386.5	23.4	537	1	P2Y8_XENLA
8	379.5	23.0	328	1	P2Y3_CHICK
9	374	22.7	374	1	P2Y2_RAT
10	373.5	22.6	373	1	P2Y2_MOUSE
11	371.5	22.5	328	1	P2Y3_MELEAGRIS
12	363.5	22.0	377	1	P2Y2_HUMAN
13	354	21.5	361	1	P2Y4_RAT
14	353	21.4	355	1	C3X1_HUMAN
15	352	21.3	365	1	P2Y4_HUMAN
16	343	20.8	361	1	P2Y4_MOUSE
17	336.5	20.4	360	1	CKR4_HUMAN
18	336	20.4	328	1	P2Y6_RAT
19	333	20.2	308	1	P2Y5_CHICK
20	332.5	20.2	354	1	C3X1_RAT
21	332	20.1	328	1	P2Y6_HUMAN
22	329	19.9	355	1	CKR1_MACMU
23	328	19.8	355	1	CKR1_HUMAN
24	326.5	19.8	346	1	CLT2_MOUSE
25	326.5	19.8	360	1	CKR4_MOUSE
26	325.5	19.7	354	1	C3X1_MOUSE
27	325.5	19.7	388	1	SSR4_HUMAN
28	319.5	19.4	345	1	CLT2_PIG
29	319	19.3	353	1	APJ_XENLA
30	318.5	19.3	359	1	IL8B_MOUSE
31	317.5	19.2	355	1	GPR1_HUMAN
32	317	19.2	368	1	SSR2_BOVIN
33	316.5	19.2	384	1	SSR4_RAT

34	315.5	19.1	355	1	GPR1_MACMU	O97664 macaca mula
35	315.5	19.1	355	1	IL8A_RABIT	P21109 oryctolagus
36	315	19.1	355	1	CKR8_HUMAN	P51685 macaca sapien
37	314	19.0	369	1	SSR2_HUMAN	P30874 homo sapien
38	314	19.0	369	1	SSR2_PIG	P34994 sus scrofa
39	314	19.0	370	1	P2Y9_HUMAN	O99677 homo sapien
40	313.5	19.0	353	1	IL8B_PANTR	Q28807 pan troglod
41	313.5	19.0	356	1	CKR8_MACMU	O97665 macaca mula
42	313.5	19.0	356	1	IL8B_CANFA	O97571 canis famil
43	313	19.0	344	1	P2Y5_HUMAN	P43657 homo sapien
44	312.5	18.9	360	1	IL8B_HUMAN	P25025 homo sapien
45	311.5	18.9	309	1	CLT2_MOUSE	Q920a1 mus musculu

ALIGNMENTS

RESULT 1					
P2YR_RAT	P2YR_RAT	STANDARD;	PRT;	373	AA.
AC	P49651;				
DT	01-FEB-1996 (Rel. 33, Created)				
DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	01-FEB-1996 (Rel. 33, Last annotation update)				
DE	P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).				
GN	P2RY1				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Insulinoma; PubMed=7779087;				
RX	MEDLINE=95298025; Hara M., Jones E.M.C., Fan Z., Bell G.I.;				
RA	Tokuyama Y., Hara M., Jones E.M.C., Fan Z., Bell G.I.;				
RT	"Cloning of rat and mouse P2Y purinoceptors."				
RL	Biochem. Biophys. Res. Commun. 211:211-218(1995).				
CC	!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS				
CC	ATP AND ADP. IN PANCREATIC ISLETS, MAY MEDIATE SOME OF THE EFFECTS				
CC	OF EXTRACELLULAR ATP ON INSULIN SECRETION.				
CC	!- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	!- TISSUE SPECIFICITY: EXPRESSED IN MUSCLE, HEART, LIVER, KIDNEY,				
CC	LUNG, BRAIN, SPLEEN, BUT NOT IN TESTIS.				
CC	!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
EMBL	U22830; AAA91303.1; .				
HSSP	P34996; 1DDD.				
DR	InterPro; IPR000276; GPCR_Rhodopsn.				
DR	Pfam; PF00001; 7tm1; 1.				
DR	PRINTS; PR00237; GPCRHHODOPSN.				
DR	PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.				
DR	PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.				
KW	G-protein coupled receptor; Transmembrane; Glycoprotein.				
FT	DOMAIN 1 52				
FT	TRANSMEM 53 74				
FT	DOMAIN 75 87				
FT	TRANSMEM 88 109				
FT	DOMAIN 110 126				
FT	TRANSMEM 127 147				
FT	DOMAIN 148 166				
FT	TRANSMEM 167 188				
FT	DOMAIN 189 218				
FT	TRANSMEM 219 238				
FT	DOMAIN 239 265				
FT	TRANSMEM 266 285				

[illegible]

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

 CC EMBL; 249205; CAA89066.1; -
 CC EMBL; U42030; AAA97873.1; -
 CC EMBL; U42029; AAA97872.1; -
 CC EMBL; S81950; AAB47091.1; -
 CC EMBL; AJ006945; CAA07339.1; -
 CC EMBL; AF018284; AAB94556.1; -
 CC HSP; P34996; 1DD0.
 CC Genew; HGNC:8539; P2RY1.
 CC MIM; 601167; -
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCRHHODOPSN.
 CC PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.
 CC PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Platelet;
 KW Blood coagulation.
 FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 53 74 1 (POTENTIAL).
 FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 88 109 2 (POTENTIAL).
 FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 127 147 3 (POTENTIAL).
 FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 167 188 4 (POTENTIAL).
 FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 219 238 5 (POTENTIAL).
 FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 266 285 6 (POTENTIAL).
 FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 304 328 7 (POTENTIAL).
 FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 124 202 BY SIMILARITY.
 FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 137 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 138 138 MISSING (IN REF. 1).
 FT SEQUENCE 373 AA; 42071 MW; 4DC7C668B4145392 CRC64;
 Query Match 28.7%; Score 474; DB 1; Length 373;
 Best Local Similarity 37.3%; Pred. No. 1.2e-24;
 Matches 107; Conservative 55; Mismatches 117; Indels 8; Gaps 6;
 QY 20 YLSAFYAIEFIFGLGNVTVFGYLFQCMKNSSNVYLPNLSISDFAFCLTLPILIKSY 79
 Db 52 YLPAVAVILVFIITGLGNSVAIMMVFHMKPSGIVSYMFNLALADFLYVLTPLIFY 111
 QY 80 ANDKG-TYGDVLCISNRVYLHNLXTLSILFLFISMDRYLLMKYPRFHFLOKKEFALLI 138
 Db 112 FNKTDIFGDMCKLQRFHNLVGLSILFLCISNHRSGVYVPLKSLGRKKNAICI 171
 QY 139 SLAVVALVTLVPLMTFFINS-VPRKESNCIDYASSGNPEHNLISGLTLLGLFLIPUS 197
 Db 172 SVLVLLVVAIVSIPILFSGTGVRRKNTICYDTTSDVLRVSYFVSMCTTVAFCVPLV 231
 QY 198 VMCFYFVYKVVFKRRSQQAALPL-DRPQLVVLAVVIFSLFPTPHYHMLNRIASRL 256
 Db 232 LILGCVGLIVRLIYKDLONS---PLRRKSYLVILVTFVFAVSIPFHVMTMNLARL 288
 QY 257 D-SWPGGCT-QKAISKSTYLTLPFLAFNSAINPIFYELMGDHYREML 301
 Db 289 DFOTPAFCAPNDRVATYQVTRGLASLNSCVDPIFLYFLAGDTFRRL 335
 RESULT 5
 ID P2YR_CHICK STANDARD; PRT; 362 AA.

AC P34996;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-FEB-1998 (Rel. 37, Last annotation update)
 DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
 GN P2RY1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93285340; PubMed=8508924;
 RA Webb T.E., Simon J., Krishak B.J., Bateson A.N., Smart T.G.,
 RT King B.F., Burnstock G., Barnard E.A.;
 "Cloning and functional expression of a brain G-protein-coupled ATP
 receptor."; FEBS Lett. 324:219-225(1993).
 RL [2]
 RN 3D-STRUCTURE MODELING.
 RP MEDLINE=97026278; PubMed=8872457;
 RX van Rhee A.M., Fischer B., van Galen P.J.M., Jacobson K.A.;
 "Modelling the P2Y purinoceptor using rhodopsin as template.";
 Drug Des. Discov. 13:133-140(1995).
 RL -1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
 ATP AND ADP. SEEMS TO MEDIATE ITS ACTION VIA A PERTUSSIS TOXIN
 INSENSITIVE G-PROTEIN, PROBABLY BELONGING TO THE GQ FAMILY THAT
 ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, GASTROINTESTINAL TRACT,
 CC SPLEEN AND LEG MUSCLE. IS NOT DETECTED IN THE HEART, LIVER,
 CC STOMACH, LUNG AND KIDNEY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC this SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

SQ SEQUENCE 362 AA; 41194 MW; A806C88FB9514761 CRC64;

Query Match 28.6%; Score 472.5; DB 1; Length 362;
 Best Local Similarity 35.9%; Pred. No. 1.4e-24;
 Matches 110; Conservative 58; Mismatches 121; Indels 17; Gaps 8;

QY 10 WLA-----TEALNK-----YYLSAFYAIEFIFGLGNVTVVFGYLFCKMKNSSNYLNF 60
 DB 22 WAAGNATTKSLKTKTGFQFYLPVTVILFVITGFLGNSVAIMFVPHMRPWSGISVYMFN 81

QY 61 LSSDFAFLCTPLILKSYANDKG-TYGDVLCISNRVYLTHTNLYTSILFTFSMDRYLL 119
 DB 82 LALADELYLVTLPALIFYFNKTDWIFGDMCKLQRFIFHVNLYGSILFTICSVHYTG 141

QY 120 MKYFFREHFLOKKEFAILISLAWALVTLVLPMLTFINS-VPKEGSNCIDYASSGNPE 178
 DB 142 VVHPLKSLGRKKNAVYVSSLVWALVAVIAPILFYSGTVGRNKITTCYDTTADAYLR 201

QY 179 HNLTPYHIMRNLIARSLD-SWPQOCT-QKAISKIYTLTRPLAFNLNSAINPIFYFLMGD 295
 DB 259 ANSVLPFHVMKTLNLRALDFTQPMCAFNDKVYATYQVTRGLASLNSCVDPILYFLAGD 318

QY 296 HYREML 301
 DB 319 TFRRL 324

RESULT 6
 P2YR_MELGA STANDARD; PRT: 362 AA.
 AC P49652;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor) (6H1
 DE orphan receptor).
 GN P2RY1.

OS Meleagris gallopavo (Common turkey).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.
 OX NCBI_TaxID=9103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94335907; PubMed=8058061;
 RA Filtz T.M., Li Q., Boyer J.L., Nicholas R.A., Harden T.K.;
 RT "Expression of a cloned P2Y purinergic receptor that couples to
 phospholipase C.";
 RL Mol. Pharmacol. 46:8-14(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97382456; PubMed=9240460;
 RA Li Q., Schachter J.B., Harden T.K., Nicholas R.A.;
 RT "The 6H1 orphan receptor, claimed to be the p2y5 receptor, does not
 mediate nucleotide-promoted second messenger responses.";
 RL Biochem. Biophys. Res. Commun. 236:455-460(1997).
 CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
 ATP AND ADP. SEEMS TO MEDIATE ITS ACTION VIA A PERTUSSIS TOXIN
 SENSITIVE G-PROTEIN, PROBABLY BELONGING TO THE GQ FAMILY THAT
 ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: MAINLY FOUND IN BLOOD, BRAIN, AND LUNG. TO A
 LESSER EXTENT IN STOMACH, GUT, AND SKELETAL MUSCLE.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U09842; AAA18784.1; -.
 DR HSSP; P34996; 1DDD.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPS.
 DR PROSITE; PS00237; G-PROTEIN RECP_F1.1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECP_F1.2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 41
 FT TRANSMEM 42 63
 FT DOMAIN 64 76
 FT TRANSMEM 77 98
 FT DOMAIN 99 115
 FT TRANSMEM 116 136
 FT DOMAIN 137 155
 FT TRANSMEM 156 177
 FT DOMAIN 178 207
 FT TRANSMEM 208 227
 FT DOMAIN 228 254
 FT TRANSMEM 255 274
 FT DOMAIN 275 292
 FT TRANSMEM 293 317
 FT DOMAIN 318 362
 FT DISULFID 113 191
 FT CARBOHYD 11 11
 FT CARBOHYD 26 26
 FT CARBOHYD 102 102
 FT CARBOHYD 186 186
 SQ SEQUENCE 362 AA; 41180 MW; 3E128AB9EB64349C CRC64;

Query Match 28.6%; Score 472.5; DB 1; Length 362;
 Best Local Similarity 35.9%; Pred. No. 1.4e-24;
 Matches 110; Conservative 58; Mismatches 121; Indels 17; Gaps 8;

QY 10 WLA-----TEALNK-----YYLSAFYAIEFIFGLGNVTVVFGYLFCKMKNSSNYLNF 60
 DB 22 WAAGNATTKSLTGTGTFQFYLPVTVILFVITGFLGNSVAIMFVPHMRPWSGISVYMFN 81

QY 61 LSSDFAFLCTPLILKSYANDKG-TYGDVLCISNRVYLTHTNLYTSILFTFSMDRYLL 119
 DB 82 LALADELYLVTLPALIFYFNKTDWIFGDMCKLQRFIFHVNLYGSILFTICSVHYTG 141

QY 120 MKYFFREHFLOKKEFAILISLAWALVTLVLPMLTFINS-VPKEGSNCIDYASSGNPE 178
 DB 142 VVHPLKSLGRKKNAVYVSSLVWALVAVIAPILFYSGTVGRNKITTCYDTTADAYLR 201

QY 179 HNLTPYHIMRNLIARSLD-SWPQOCT-QKAISKIYTLTRPLAFNLNSAINPIFYFLMGD 295
 DB 259 ANSVLPFHVMKTLNLRALDFTQPMCAFNDKVYATYQVTRGLASLNSCVDPILYFLAGD 318

QY 296 HYREML 301
 DB 319 TFRRL 324

RESULT 7
 P2Y8_XENLA STANDARD; PRT: 537 AA.
 ID P2Y8_XENLA
 AC P79928;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE P2Y purinoceptor 8 (P2Y8).
 GN P2RY8.

OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Neural plate;
 RX MEDLINE=97284734; PubMed=9139711;
 RA Bogdanov Y.D., Dale L., King B.F., Whittock N., Burnstock G.;
 RT "Early expression of a novel nucleotide receptor in the neural plate
 of Xenopus embryos";
 RL J. Biol. Chem. 272:12583-12590(1997).
 CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ATP, UTP, CTP, GTP AND ITP.
 CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
 CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC MAY PLAY A KEY ROLE IN THE EARLY DEVELOPMENT OF NEURAL TISSUE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X99953; CAA68213.1; .
 DR HSP; P34996; 1DDD.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; 1.
 DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_F2_1; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 49
 FT TRANSMEM 50 70
 FT DOMAIN 71 79
 FT TRANSMEM 80 100
 FT DOMAIN 101 118
 FT TRANSMEM 119 139
 FT DOMAIN 140 161
 FT TRANSMEM 162 182
 FT DOMAIN 183 210
 FT TRANSMEM 211 231
 FT DOMAIN 232 254
 FT TRANSMEM 255 275
 FT DOMAIN 276 292
 FT TRANSMEM 293 316
 FT DOMAIN 317 337
 FT DISULFID 116 193
 FT CARBOHYD 26 26
 FT CARBOHYD 29 29
 FT SEQUENCE 537 AA; 6204 MW; B2CF24812F3C19F2 CRC64;
 Query Match 23.4%; Score 386.5; DB 1; Length 537;
 Best Local Similarity 35.3%; Pred. No. 9.2e-19;
 Matches 104; Conservative 51; Mismatches 125; Indels 15; Gaps 7;
 QY 22 LSAPYATEFTFGLGNVTYVGVFLCMKNWSSNVFLNLSISDFALCTPLILIKSYAN 81
 DB 46 LPVSIVAFVWGLPNTAAMWIFIAKMRPNWPTVYVFNALSDLTLYLSLPTLVYYA- 104
 QY 82 DKGT--YGDVLCISNRVLTHTNLTSTLFTFISMDRYLMKYPFFREHFLQKEFAILIS 139
 DB 105 DKNWNPGEVLCVLRFLFYANLYSSILFTLCISVHRYGVCHPITSLRMMNAKHAYIC 164
 QY 140 LAVWALTVLEVPMLTINSVPKEGNCIDYASSGNPEHNLISCLTLGLPLISVYM 199
 DB 165 ALVWLSVTLCVLPNLTFTVSPKVKNTICHDTRPEDFARYVEYSTAIMCLLFGICLLII 224
 QY 200 CFFYKHWVFLKRR--SQQQATALPLDKPQ--RLVVLAVVIFSIETPTPHMNLRIASR 255

Db 225 AGCYGLMRELKPIVSGNQOT-LPSYKRSIKTIIFVMAFAICFMPPHITRTYYAR 283
 QY 256 LDSWPGCQTQAKSI---YTLTRPLAFNSAINPIFYFLMGDHYREMLISKFRQ 307
 Db 284 L-----LGIKYALNINVITVKVTRPLASANSCLDPILYFLANDRYRRLIRVRR 334
 [1]
 RESULT 8
 P2Y3_CHICK
 ID P2Y3_CHICK STANDARD; PRT; 328 AA.
 AC Q98907;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE P2Y purinoceptor 3 (P2Y3) (Nucleoside diphosphate receptor).
 GN P2Y3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96319774; PubMed=8700132;
 RA Bateson A.N., Burnstock G., Barnard E.A.;
 RA "A novel G protein-coupled P2 purinoceptor (P2Y3) activated
 preferentially by nucleoside diphosphates";
 RT Mol. Pharmacol. 50:258-265(1996).
 RL FUNCTION: RECEPTOR FOR EXTRACELLULAR ADP > UTP > ATP = UDP. THE
 CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X98283; CAA66930.1; .
 DR HSP; P34996; 1DDD.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; FALSE_NEG.
 DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_F2_1; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 22
 FT TRANSMEM 23 43
 FT DOMAIN 44 57
 FT TRANSMEM 58 78
 FT DOMAIN 79 96
 FT TRANSMEM 97 117
 FT DOMAIN 118 139
 FT TRANSMEM 140 160
 FT DOMAIN 161 189
 FT TRANSMEM 190 210
 FT DOMAIN 211 231
 FT TRANSMEM 232 252
 FT DOMAIN 253 275
 FT TRANSMEM 276 298
 FT DOMAIN 299 323
 FT CARBOHYD 5 5
 FT DISULFID 94 172
 FT SEQUENCE 328 AA; 37586 MW; 7A3BF1C91F54FAAB CRC64;
 Query Match 23.0%; Score 379.5; DB 1; Length 328;

[illegible]

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92281707; PubMed=7685114;
RA Lustig K.D., Shiao A.K., Brake A.J., Julius D.;
RT "Expression cloning of an ATP receptor from mouse neuroblastoma
cells";
RL Proc. Natl. Acad. Sci. U.S.A. 90:5113-5117(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96316177; PubMed=8739319;
RA Enomoto K., Furiya K., Moore R.C., Yamagishi S., Oka T., Maeno T.;
RT "Expression cloning and signal transduction pathway of P2U receptor
in mammary tumor cells";
RL Biol. Signals 5:9-21(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Head, and Liver;
RC MEDLINE=21083660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arawaka T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamahata I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baidarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP MUTAGENESIS OF LYS-107; ARG-110; HIS-262; ARG-265; LYS-289 AND
RP ARG-292.
RX MEDLINE=95181393; PubMed=7876172;
RA Erb L., Garrad R.C., Wang Y., Quinn T., Turner J.T., Weisman G.A.;
RT "Site-directed mutagenesis of P2U purinoceptors. Positively charged
amino acids in transmembrane helices 6 and 7 affect agonist potency
and specificity";
RL J. Biol. Chem. 270:4185-4188(1995).
CC [-] FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
CC activate a phosphatidylinositol-calcium second messenger system.
CC The affinity range is UTP = ATP > ATP-gamma-S >> 2-methylthio-ATP
CC = ADP.
CC [-] SUBCELLULAR LOCATION: Integral membrane protein.
CC [-] TISSUE SPECIFICITY: SPLEEN, TESTIS, KIDNEY, LIVER, LUNG, HEART AND
CC BRAIN.
CC [-] SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L14751; AAA39871.1; -;
CC EMBL; S83099; AAB50735.1; -;
CC EMBL; AK005013; BAB23746.1; -;

DR EMBL; AK017378; BAB30719.1; -;
DR EMBL; BC005613; AAB06613.1; -;
DR HSSP; P34996; 1DDD.
DR MGD; MGI:105107; P2ry2.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR_F1_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein.
KW DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 33 59 1 (POTENTIAL).
FT DOMAIN 60 70 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 71 93 2 (POTENTIAL).
FT DOMAIN 94 110 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 111 129 3 (POTENTIAL).
FT DOMAIN 130 152 4 (POTENTIAL).
FT TRANSMEM 153 172 5 (POTENTIAL).
FT DOMAIN 173 194 6 (POTENTIAL).
FT TRANSMEM 195 220 7 (POTENTIAL).
FT DOMAIN 221 246 8 (POTENTIAL).
FT TRANSMEM 247 269 9 (POTENTIAL).
FT DOMAIN 270 287 10 (POTENTIAL).
FT TRANSMEM 288 309 11 (POTENTIAL).
FT DOMAIN 310 373 12 (POTENTIAL).
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT BY SIMILARITY.
FT DISULFID 106 183 K->I: NO EFFECT ON RECEPTOR ACTIVATION.
FT MUTAGEN 107 107 R->L: NO EFFECT ON RECEPTOR ACTIVATION.
FT MUTAGEN 110 110 H->L: DECREASE IN RECEPTOR ACTIVATION.
FT MUTAGEN 262 262 R->L: DECREASE IN RECEPTOR ACTIVATION.
FT MUTAGEN 265 265 K->I: NO EFFECT ON RECEPTOR ACTIVATION.
FT MUTAGEN 289 289 R->L: DECREASE IN RECEPTOR ACTIVATION.
FT MUTAGEN 292 292 E -> D (IN REF. 2).
FT CONFLICT 17 17 S -> R (IN REF. 2).
FT CONFLICT 120 120 T -> N (IN REF. 2).
FT CONFLICT 125 125 V -> M (IN REF. 2).
FT CONFLICT 196 196 V -> L (IN REF. 3).
FT CONFLICT 263 263 D -> N (IN REF. 2).
FT CONFLICT 355 355 KDI -> PYV (IN REF. 2).
FT CONFLICT 369 371
SQ SEQUENCE 373 AA; 42174 MW; 590BBE502E41B3AE CRC64;

Query Match 22.6%; Score 373.5; DB 1; Length 373;
Best Local Similarity 33.5%; Pred No 4.7e-18;
Matches 94; Conservative 50; Mismatches 130; Indels 7; Gaps 5;

QY 19 KY-YLSAFYAEIEIFGLGNVTVVGYLFCKMKNSSNYLFNLSISDFAPLCTPLIK 77
DB 32 KYVLLPVSYGVVCLNVALYIFLCRLKTNASTYMFHLAVSDSLYASLPV 91
QY 78 SYA-NDKGTGVDVLCISNRYVHLNLTSLFTFISMDRYLLMKYPFREHFLQKEFAI 136
DB 92 YYARGDHPFSTVLCKLVRFLEYTNLYCSILFLTCISVHRCGLVRLPLHSRNGRYAR 151
QY 137 LISLAVWLVLEVLPMLETFNSVPKESGNCIDYASSGNPEHNLISLCLTGLGLPL 196
DB 152 RVAAYVWLVLAQAPVLYFVITTSVRGTRTCHTSARELSEHFVAYSSVMGLLPAVPE 211
QY 197 SVMCFYKVVVFKRRSQQAATPLDKPQ-RLVWLVVIFSLTTPHIMNLRIAS 254
DB 212 SVILCYVLMARELLKPAYGTGGPLPRAKRKSVRTIALVAVFALCFEPFHVTRTLAYSF 271
QY 255 RLDSWPOQC-TQAKTSYITITRPLAFUNSAINFIFLMG 294
DB 273 R-SLDLSCHTLNAINMAYKITRPLASANSCLDPVLYFLAG 310

RESULT 11
ID P2Y3_MELGA STANDARD; PRT; 328 AA.
AC O93361;
DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE P2Y purinoceptor 3 (P2Y3) (Nucleoside diphosphate receptor).
 GN P2Y3.
 OS Meleagris gallopavo (Common turkey).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.
 OX NCBI_TaxID=9103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98401046; PubMed=9730913;
 RA Li Q., Olesky M., Palmer R.K., Harden T.K., Nicholas R.A.;
 RA "Evidence that the p2y3 receptor is the avian homologue of the
 RT mammalian P2Y6 receptor.";
 RL Mol. Pharmacol. 54:541-546(1998).
 CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR UDP> ADP = UTP. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
 CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF069555; AAC23863.1; -.
 DR HSP; P34996; 1DD0.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; FALSE_NEG.
 DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 22 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 23 43 1 (POTENTIAL).
 FT DOMAIN 44 57 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 58 78 2 (POTENTIAL).
 FT DOMAIN 79 96 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 97 117 3 (POTENTIAL).
 FT DOMAIN 118 139 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 140 160 4 (POTENTIAL).
 FT DOMAIN 161 189 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 190 210 5 (POTENTIAL).
 FT DOMAIN 211 231 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 232 252 6 (POTENTIAL).
 FT DOMAIN 253 275 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 276 298 7 (POTENTIAL).
 FT DOMAIN 299 323 EXTRACELLULAR (POTENTIAL).
 FT CARBOHYD 5 5 N-LINKED (GLCNAC...) (POTENTIAL).
 FT DISULFID 94 172 BY SIMILARITY.
 SQ SEQUENCE 328 AA; 37594 MW; B74D49B99C716A5 CRC64;
 Query Match 22.5%; Score 371.5; DB 1; Length 328;
 Best Local Similarity 29.4%; Pred. No. 5.7e-18;
 Matches 87; Conservative 61; Mismatches 137; Indels 11; Gaps 4;
 QY 22 LSAYFAIEFIFGLGNVTWVFGYLFMCNKNSSNVYLFNLSISDFALCFLPLILIKSYA- 80
 DB 24 LPLVTSVVFLLGLPLNAVVGIIQWLARKALTRTIYMLNATADLLYVCSLPLIINYTQ 83
 QY 81 NDKGTGVDVLCISNRYLVHNNLYTSILFLTFISMDRYLLMKYPFRE-HFLQKEFAILIS 139
 DB 84 KDYPFGDFCTCKFRFQFYFTNLHGSILFLCISVQRYMGICHPLASWKKKKLTLWVC 143
 QY 140 LAVNALVTLVPLMLTFINSVPKEEGSNCTDYASSGNPHNLIYSLCLLLGLFLPLSVYM 199
 DB 144 AAVWFIVIAOCLPTFFVASTGTQNRITVCYDLSPPDORSASYFPYGITLITGFLFPFAI 203
 QY 200 CFFYKVMVFLKRRSQOQATAL--PLDKPQRLVVLAVVIFSLITPTPHIMRNLRITASRD 257

DB 204 LACYCSMARILCQDELIGLAVHKKKAVRMIIIVIVFSISFPFHLTKIYLIIVRS 263
 QY 258 SWPGCTQKAISYITLTPPLAFNLNSAINPFIYFLMGDHYRE-----MLISKFR 306
 DB 264 PTLPCPTLQAFATAYKCTRFPSMNSVLDPIFYFTQRFKRESTRYLDDKMSKWR 319
 RESULT 12
 P2Y2_HUMAN
 ID P2Y2_HUMAN STANDARD; PRT; 377 AA.
 AC P41231; Q96DM8;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)
 DE (Purinoceptor receptor).
 GN P2RY2 OR P2RU1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Airway epithelium;
 RX MEDLINE=94211846; PubMed=8159738;
 RA Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H.,
 RA Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;
 RA "Cloning and expression of a human P2U nucleotide receptor, a target
 RT for cystic fibrosis pharmacotherapy.";
 RT Proc. Natl. Acad. Sci. U.S.A. 91:3275-3279(1994).
 RN [2]
 RP REVISIONS.
 RX MEDLINE=95108098; PubMed=7809171;
 RA Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H.,
 RA Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;
 RT "Cloning and expression of a human P2U nucleotide receptor, a target
 RT for cystic fibrosis pharmacotherapy.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:13067-13067(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney, and Leukocyte;
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
 CC activate a phosphatidylinositol-calcium second messenger system.
 CC The affinity range is UTP = ATP > ATP-gamma-S >> 2-methylthio-ATP
 CC = ADP.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: SPLEEN, KIDNEY, LIVER, LUNG, HEART AND
 CC BRAIN.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U07225; AAC04923.1; -.
 DR EMBL; BC012104; AAH12104.1; -.
 DR EMBL; BC028135; AAH28135.1; -.
 DR PIR; A54946; A54946.
 DR HSP; P34996; 1DD0.
 DR Genew; HGNC:8541; P2RY2.
 DR MIM; 600041; -.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.

G-protein coupled receptor; Transmembrane; Glycoprotein.
KW DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 33 59 1 (POTENTIAL).
FT TRANSSEM 60 70 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 71 93 2 (POTENTIAL).
FT TRANSSEM 94 110 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 111 129 3 (POTENTIAL).
FT TRANSSEM 130 152 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 153 172 4 (POTENTIAL).
FT TRANSSEM 173 194 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 195 220 5 (POTENTIAL).
FT TRANSSEM 221 246 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 247 269 6 (POTENTIAL).
FT TRANSSEM 270 287 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 288 309 7 (POTENTIAL).
FT TRANSSEM 310 377 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 9 9 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 13 13 BY SIMILARITY.
FT CONFLICT 106 193 R -> S (IN REF. 3; AAH12104).
FT CONFLICT 312 342 E -> G (IN REF. 1).
FT CONFLICT 350 350 S -> F (IN REF. 1).
FT CONFLICT 359 359 S -> F (IN REF. 1).
SQ SEQUENCE 377 AA; 42289 MW; EE557A857A269AC6 CRC64;
Query Match 22.0%; Score 363.5; DB 1; Length 377;
Best Local Similarity 33.5%; Pred. No. 2.2e-17;
Matches 94; Conservative 52; Mismatches 128; Indels 7; Gaps 5;
QY 19 KY-YLSAFYAEIFGLGNVTVFVGYFCMKNNSSNVYFNLSISDFAFCTLPILIK 77
DB 32 KYVLLPSYGVVGLCLNAVALYIFCLRLKTNASTTMYFHLAVSDALYASLPVLY 91
QY 78 SYA-NDKGTGVDLCISNRVLTHTNLTSLFTFTSMORYLMLKYPFEHFLQKKEAI 136
DB 92 YYAGRDHWPSTVLCKLVRLFTNYCYSLFTLCISVHRGVLRLPLSLKRWGRARYAR 151
QY 137 LISLAVNALVTLEVPMLTFINSVPKEEGSNCDYASSGNPEHNLISCLTLGLFLPL 196
DB 152 RVAGVWVVLACQAPVLYEFTVSARGRVTCHTSAPELFSRFVAYSSVMLGLFAVP 211
QY 197 SVMGFFYYKMWVFLKRSSQQAATLPDKPQRLVAVV--IFSILFTYHIMNRLIAS 254
DB 212 AVILCVLVMARLLKPAYGTSGLPRAKRKSVRTAVLAVLAVFALCFPLPHVTRTYYSF 271
QY 255 RLDSWPGC-TQAKSIYTLTLPFLNAINPIFYFLMG 294
DB 272 R--SLDLSCHTLNAINMAYKVTREPLASANSCLDPVLYFLAG 310
RESULT 13
ID P2Y4_RAT STANDARD; PRT; 361 AA.
AC O35811;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 4 (P2Y4).
GN P2Y4 OR P2Y4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Liver;
RA Bogdanov Y.D., Wildman S., King B.F., Burnstock G.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=98421785; PubMed=9751165;
RA Webb T.E., Henderson D., Roberts J.A., Barnard E.A.;
RT "Molecular cloning and characterization of the rat P2Y4 receptor.";

J. Neurochem. 71:1424-1434 (1998).
-|- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that activate a phosphatidylinositol-calcium second messenger system.
CC Not activated by ADP or UDP.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- TISSUE SPECIFICITY: Widely expressed at low levels. In brain, higher expression in the pineal gland and ventricular system.
CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; Y14705; CAA75007.1; --
CC EMBL; Y11433; CAA72241.1; --
CC HSP; P34996; IDDD.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHHODOPS.
CC PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECEP_FL_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein.
KW DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 31 58 1 (POTENTIAL).
FT TRANSSEM 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 69 91 2 (POTENTIAL).
FT TRANSSEM 92 108 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 109 127 3 (POTENTIAL).
FT TRANSSEM 128 149 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 150 170 4 (POTENTIAL).
FT TRANSSEM 171 192 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 193 218 5 (POTENTIAL).
FT TRANSSEM 219 242 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 243 265 6 (POTENTIAL).
FT TRANSSEM 266 283 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 284 305 7 (POTENTIAL).
FT TRANSSEM 306 361 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 104 181 BY SIMILARITY.
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 361 AA; 40893 MW; 0377F96E54B449A3 CRC64;
Query Match 21.5%; Score 354; DB 1; Length 361;
Best Local Similarity 30.6%; Pred. No. 8.7e-17;
Matches 86; Conservative 60; Mismatches 125; Indels 10; Gaps 5;
QY 26 YAEIFEGLLGNVTVFVGYFCMKNNSSNVYFNLSISDFAFCTLPILIKSY-ANDKG 84
DB 38 YAVVFLGLNALNAPTILWLFELRPRWDATATYMFHALSDTLVLSLPTLVVYAAARNHW 97
QY 85 TYGDVLCISNRVLTHTNLTSLFTFTSMORYLMLKYPFEHFLQKKEFALLISLAYWA 144
DB 98 PFTGLCKEVRFLFYWNLYCSVLFTCTISVHRYLGHICHPRAIRWGRPRFASLLCGLGWL 157
QY 145 LYTELVPLMLTFINSVPKEEGSNCDYASSGNPEHNLISCLTLGLFLPLSVNCFY 204
DB 158 VVAGCLVPLNFFVTNANTGTTILCHDTTLPEFDHYVVFSSAVVLLFGLPELITLVCYG 217
QY 205 KMWVFLKR---RSQQAATLPDKPQRLVAVLAVIFSILFTYHIMNRLIASRLDSWPQ 261
DB 218 LMARRLYRPLPGAGQSSRL---RSLRTAVVLTFAVCFVFPFHTITRYIYQARL--LQA 272
QY 262 GC-TQAKSIYTLTLPFLNAINPIFYFLMGDHYREML 301
DB 273 DCHVLNIVNVYXVTRPLASANSCLDPVLYLFTGDKYRNQL 313
RESULT 14
ID C3X1_HUMAN STANDARD; PRT; 355 AA.
C3X1_HUMAN

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 15:16:42 ; Search time 42 Seconds
(without alignments)
1555.165 Million cell updates/sec

Title: US-09-891-138a-2
Perfect score: 1650
Sequence: 1 MAONLSCENWLATEAILNKY.....REMLISKFRQYFKSLTSFRT 317

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: sp.archaea.*
 - 2: sp.bacteria.*
 - 3: sp.fungi.*
 - 4: sp.human.*
 - 5: sp.invertebrate.*
 - 6: sp.mammal.*
 - 7: sp.mhc.*
 - 8: sp.organelle.*
 - 9: sp.phage.*
 - 10: sp.plant.*
 - 11: sp.rodent.*
 - 12: sp.virus.*
 - 13: sp.vertibrate.*
 - 14: sp.unclassified.*
 - 15: sp.virus.*
 - 16: sp.bacteriap.*
 - 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1634	99.0	317	11 Q99MT6	Q99mt6 mus musculus
2	1238.5	75.1	330	4 Q9BXA5	Q9bxa5 homo sapien
3	1238.5	75.1	334	4 Q8TDQ8	Q8tdq8 homo sapien
4	482.5	29.2	361	13 Q90X57	Q90x57 xenopus lae
5	461	27.9	357	13 Q9DE05	Q9de05 raja erinac
6	426	25.8	337	4 Q96P68	Q96p68 homo sapien
7	422.5	25.6	374	13 Q57466	Q57466 meleagris g
8	353	21.4	361	11 Q8VHP3	Q8vhp3 cavia porce
9	338	20.5	328	11 Q9ERK9	Q9erh9 mus musculus
10	335	20.3	355	6 Q9MTJ8	Q9myj8 callithrix
11	333	20.2	355	11 Q9JLY8	Q9jly8 rattus norv
12	326	19.8	351	6 Q9MTJ9	Q9myj9 oryctolagus
13	320	19.4	360	11 Q9IZH4	Q9izh4 rattus norv
14	318.5	19.3	309	11 Q8RS28	Q8rs28 mus musculus
15	316.5	19.2	355	6 Q95LH1	Q95lh1 macaca fasc
16	316	19.2	355	11 Q91VP9	Q91vp9 mus musculus

17	316	19.2	367	13 Q9PVG0	Q9pvq0 carassius a
18	315	19.1	355	4 Q9BYX5	Q9byx5 homo sapien
19	314	19.0	356	4 Q96TF2	Q96tf2 homo sapien
20	313	19.0	367	13 Q9PVF9	Q9pvf9 carassius a
21	312.5	18.9	346	4 Q9GEO0	Q9ge00 homo sapien
22	312.5	18.9	390	13 Q8QGO4	Q8qgq4 carassius a
23	308	18.7	346	11 Q91Y73	Q91y73 mus musculus
24	306.5	18.6	392	6 Q9BDQ4	Q9bdq4 canis famil
25	306	18.5	337	11 Q921N3	Q92ln3 mus musculus
26	301	18.2	380	13 Q9DGO6	Q9dgg6 carassius a
27	300	18.2	354	12 Q91OT6	Q91qt6 human cytom
28	299	18.1	346	6 Q95KS6	Q95ks6 ovis aries
29	299	18.1	354	12 Q91P65	Q91pe5 human cytom
30	298	18.1	354	12 Q91Q00	Q91qu0 human cytom
31	298	18.1	354	12 Q91OK7	Q91ok7 human cytom
32	298	18.1	363	4 Q8TDS4	Q8tds4 homo sapien
33	297.5	18.0	358	6 Q9NOM0	Q9nom0 ovis aries
34	296	17.9	346	4 Q9BXC0	Q9bxc0 homo sapien
35	296	17.9	354	12 Q91P67	Q91pe7 human cytom
36	296	17.9	354	12 Q91P66	Q91pe6 human cytom
37	295	17.9	354	12 Q91P68	Q91pe8 human cytom
38	295	17.9	354	12 Q91P69	Q91pe9 human cytom
39	295	17.9	354	12 Q91QT9	Q91qt9 human cytom
40	295	17.9	354	12 Q91QT8	Q91qt8 human cytom
41	295	17.9	354	12 Q91QT7	Q91qt7 human cytom
42	295	17.9	354	12 Q91OH3	Q91oh3 human cytom
43	295	17.9	354	12 Q91OE0	Q91oe0 human cytom
44	294.5	17.8	360	13 Q8QPR6	Q8qpr6 xenopus lae
45	294	17.8	354	12 Q91QU2	Q91qu2 human cytom

ALIGNMENTS

RESULT 1

Q99MT6 ID Q99MT6 PRELIMINARY; PRT; 317 AA.
AC Q99MT6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE G-protein coupled receptor GPR91.
GN GPR91.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL;
RX MEDLINE=21172992; PubMed=11273702;
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in
the discovery of new g-protein coupled receptors.";
RL J. Mol. Biol. 307:799-813(2001).
DR EMBL; AF295367; AAK01867.1; -
DR HSSP; P34996; 1DDO.
DR MGD; MGI:1934135; Gpr91.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE. 317 AA; 36701 MW; 4E22F0608F928E36 CRC64;

Query Match 99.0%; Score 1634; DB 11; Length 317;
Best Local Similarity 99.1%; Pred. No. 5.8e-138;
Matches 314; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAONLSCENWLATEAILNKYLSAFYAIFIFGLLGNVTVFYGLFCMKNNSSNVYLFN 60
|||||
Db 1 MAONLSCENWLATEAILNKYLSAFYAIFIFGLLGNVTVFYGLFCMKNNSSNVYLFN 60
|||||

```

QY 61 LSISDFALCTLPILIKSYANDKGYGVDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 120
Db 61 LSISDFALCTLPILIKSYANDKGYGVDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 120
QY 121 KYPRFHFLOKKEFAILLISLAWALVTLEVLPMLTFFINSVPKEGSCNIDYASSGNPEHN 180
Db 121 KYPRFHFLOKKEFAILLISLAWALVTLEVLPMLTFFINSVPKEGSCNIDYASSGNPEHN 180
QY 181 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATPLDKPQRLVVLAVVIFSIL 240
Db 181 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATPLDKPQRLVVLAVVIFSIL 240
QY 241 FPTYHIMRNLRITASRLDSWPQCTQKAIKSIYTLRPLAFINSAINPIFYFLMGDHYRE 300
Db 241 FPTYHIMRNLRITASRLDSWPQCTQKAIKSIYTLRPLAFINSAINPIFYFLMGDHYRE 300
QY 301 LISKPROYFKSLTSF 317
Db 301 LISKPROYFKSLTSF 317

RESULT 2
Q9BXA5 PRELIMINARY; PRT; 330 AA.
AC Q9BXA5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE G-protein coupled receptor 91.
GN GPR91.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP MEDLINE-21172992; PubMed-11273702;
RX Wittenberger T., Schaller H.C., Hellebrand S.;
RA "An expressed sequence tag (est) data mining strategy succeeding in
RT the discovery of new g-protein coupled receptors.";
RL J. Mol. Biol. 307:799-813(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF348078; AAK29080.1; -.
DR HSP; P34996; 1DD0.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1. 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 330 AA; 38283 MW; 6D3C70E340866BAB CRC64;

```

```

Query Match 75.1%; Score 1238.5; DB 4; Length 330;
Best Local Similarity 72.2%; Pred. No. 1.2e-102;
Matches 228; Conservative 42; Mismatches 45; Indels 1; Gaps 1;

QY 1 MAQNLSNENLATEAILNKYLSAFYAIEFIFGLGNVTVVFGYLCMKNNSSNYLFN 60
Db 1 MAWNATCKNMLAAEALEKYLSIFGYIEFVGVGLNTIVVGYIFSLKNNSSNYLFN 60
QY 61 LSISDFALCTLPILIKSYANDKGYGVDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 120
Db 61 LSISDFALCTLPILIKSYANDKGYGVDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 120
QY 121 KYPRFHFLOKKEFAILLISLAWALVTLEVLPMLTFFINSVPKEGSCNIDYASSGNPEHN 180
Db 121 KYPRFHFLOKKEFAILLISLAWALVTLEVLPMLTFFINSVPKEGSCNIDYASSGNPEHN 180
QY 181 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATPLDKPQRLVVLAVVIFSIL 240
Db 181 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATPLDKPQRLVVLAVVIFSIL 240

```

```

QY 241 FPTYHIMRNLRITASRLDSWPQCTQKAIKSIYTLRPLAFINSAINPIFYFLMGDHYRE 299
Db 241 FPTYHIMRNLRITASRLDSWPQCTQKAIKSIYTLRPLAFINSAINPIFYFLMGDHYRE 299
QY 300 MLISKPROYFKSLTSF 315
Db 300 MLMNQLRHNFKSLTSF 316

RESULT 3
Q8TDQ8 PRELIMINARY; PRT; 334 AA.
AC Q8TDQ8;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE P2Y purinoceptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP Zhang W., Li N., Wan T., Cao X.;
RA "Human P2Y purinoceptor 1";
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF247785; AAL95690.1; -.
SQ SEQUENCE 334 AA; 38697 MW; 33146E1AD87F0E81 CRC64;

Query Match 75.1%; Score 1238.5; DB 4; Length 334;
Best Local Similarity 72.2%; Pred. No. 1.2e-102;
Matches 228; Conservative 42; Mismatches 45; Indels 1; Gaps 1;

QY 1 MAQNLSNENLATEAILNKYLSAFYAIEFIFGLGNVTVVFGYLCMKNNSSNYLFN 60
Db 5 MAWNATCKNMLAAEALEKYLSIFGYIEFVGVGLNTIVVGYIFSLKNNSSNYLFN 64
QY 61 LSISDFALCTLPILIKSYANDKGYGVDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 120
Db 65 LSISDFALCTLPILIKSYANDKGYGVDVLCISNRYVLTHTNLTYSILFLTFISMDRYLLM 124
QY 121 KYPRFHFLOKKEFAILLISLAWALVTLEVLPMLTFFINSVPKEGSCNIDYASSGNPEHN 180
Db 125 KYPRFHFLOKKEFAILLISLAWALVTLEVLPMLTFFINSVPKEGSCNIDYASSGNPEHN 184
QY 181 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATPLDKPQRLVVLAVVIFSIL 240
Db 185 LIYSCLTLGLFLPLSVCMCFYKMWVFLKRRSQOQATPLDKPQRLVVLAVVIFSIL 244
QY 241 FPTYHIMRNLRITASRLDSWPQCTQKAIKSIYTLRPLAFINSAINPIFYFLMGDHYRE 299
Db 245 FPTYHIMRNLRITASRLDSWPQCTQKAIKSIYTLRPLAFINSAINPIFYFLMGDHYRE 304
QY 300 MLISKPROYFKSLTSF 315
Db 305 MLMNQLRHNFKSLTSF 320

RESULT 4
Q90X57 PRELIMINARY; PRT; 361 AA.
ID Q90X57;
AC Q90X57;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE P2Y nucleotide receptor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
[1]

```

Query Match 27.9%; Score 461; DB 13; Length 357;

```
QY      || :      :|| :| | : || : || || | : || :
// KSYAN-DRGTYGDVLCISNRIVLHTNLVTSLEFLTFTISMDRYE
```

Db	91	HYASGENWIFGDFMCKFIRESFHNLVSSILFTCTSFYFCVLIHPMSCFSIHKTRCA	150
Qy	136	ILISLAVNALVTLEVLPLMTLINSVPKEEGSNCIDYASSGNPEHNL--YSLCLTLGLFL	193
Db	151	VVACAVWIIISIVAPITPMTLITSTNRNSACDLTSS--DELNTIKWNLILTAATTC	208
Qy	194	IPLSVWCFFYKVMVFLKRRSQOQATPLDKPQRLVVLAVVIFSLFTPHIMNLRJA	253
Db	209	LPLVITVLCYTTIIHTHGLQDSCL--KQKARRLTIILLLAFYVCFPLPHILIRIE	266
Qy	254	SRLDSPGOGCT-OKAIKSIYTLTRPLAFNLAINPIFYFLMGDHYREMLISKFR	306
Db	267	SKLLS--ISCSIEHQIHEAYIVSRPLAALNTFGLNLLVYVSDNFQQAQVCTVR	318
RESULT 7			
ID	O57466	PRELIMINARY;	PRT; 374 AA.
AC	O57466;		
DT	01-JUN-1998 (TReMBLrel. 06, Created)		
DT	01-JUN-1998 (TReMBLrel. 06, Last sequence update)		
DT	01-DEC-2001 (TReMBLrel. 19, Last annotation update)		
DE	G protein coupled P2Y nucleotide receptor.		
GN	TP2Y.		
OS.	Meleagris gallopavo (Common turkey).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.		
OX	NCBI_TaxID=9103;		
RN	[1]		
RF*	SEQUENCE FROM N.A.		
RC	TISSUE=BLOOD;		
RX	MEDLINE=98086419; PubMed=9415702;		
RA	Boyer J.L., Waldo G.L., Harden T.K.;		
RT	"Molecular cloning and expression of an avian G protein-coupled P2Y receptor";		
RL	Mol. Pharmacol. 52:928-934(1997).		
DR	EMBL; AF031897; AAC60339.1; -		
DR	HSSP; P34996; 1DDO.		
DR	InterPro: IPR000276; GPCR_Rhodopsn.		
DR	Pfam: PF00001; 7tm.1.1		
DR	PRINTS; PRO0237; GPCR_RHODOPSN		
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.		
DR	PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.		
KW	Receptor.		
SQ	SEQUENCE 374 AA; 42594 MW; 849C465722BD02B CRC64;		
Query Match 25.6%; Score 422.5; DB 13; Length 374;			
Best Local Similarity 33.7%; Pred. No. 9.4e-30;			
Matches 106; Conservative 59; Mismatches 119; Indels 31; Gaps 8;			
Qy	10	WL-----APEA--ILNKY---LSAFYAIEFIFGLGNVTVVFGYLCMKNNSSNVY	57
Db	20	WLGNTTAAAEAKCVNEEFKILLPISYGIYVVGVLPLNSWAMWIFVSRMRPNATTTY	79
Qy	58	LFNLISDPAFLCTLPILIKSYAN-DKGYGDVLCISNRVYVLTNLTSTILFTFISMDR	116
Db	80	MFNLASDITLVFSLTVYVYADRNWPFKGVKFCIVRFLFYANLYSSILFTCLISVHR	139
Qy	117	YLLMKYFPEHFLQKKEFALLISLAVNALVTLEVLPLMTLINSVPKEEGSNCIDYASSGN	176
Db	140	YMGICHPIRSLKWKTKHARLICGVNLVYVICLIENLIFVTSSKDNSTLCHDTTKPEE	199
Qy	177	PEHNLVSLCLTLGLFPLISVMCFYKVMVFLKRRSQOQATPLDKPQ-----	227
Db	200	FDHVHVYSSSMALLFGIPFLVIVWCYCLMAKRLCKRS-----FPSSPRPVPSYKRSI	253
Qy	228	RLVVLAVVIFSLFTPHIMNLRIRASRLDSWFGQC-TQKAISYITLTPPLAFNLAIN	286
Db	254	KMIITVTFVFAICVPVPHITRTLYTSRY--FOADCOTLIINFTYKTRPLASINCLD	311
Qy	287	PIFYFLMGDHYREML	301
Db	312	PILYFMAGDKYRGEL	326
RESULT 8			
ID	Q8VHP3	PRELIMINARY;	PRT; 361 AA.
AC	Q8VHP3;		
DT	01-MAR-2002 (TReMBLrel. 20, Created)		
DT	01-MAR-2002 (TReMBLrel. 20, Last sequence update)		
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)		
DE	CC-chemokine receptor 4.		
OS	Cavia porcellus (Guinea pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.		
OX	NCBI_TaxID=10141;		
RN	[1]		
RF*	SEQUENCE FROM N.A.		
RA	Jopling L.A., Sabroe I., Andrew D.P., Mitchell T.J., Li Y.,		
RA	Hodge M.R., Williams T.J., Pease J.E.;		
RT	"The identification, characterization and distribution of guinea pig CCR4 and epitope mapping of a blocking antibody.";		
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF431971; AAL57488.1; -		
DR	InterPro: IPR000276; GPCR_Rhodopsn.		
DR	Pfam: PF00001; 7tm.1.1		
DR	PRINTS; PRO0237; GPCR_RHODOPSN		
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.		
DR	PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.		
KW	Receptor.		
SQ	SEQUENCE 361 AA; 41064 MW; 9304E897D4FD6839 CRC64;		
Query Match 21.4%; Score 353; DB 11; Length 361;			
Best Local Similarity 30.0%; Pred. No. 1.5e-23;			
Matches 86; Conservative 62; Mismatches 131; Indels 8; Gaps 5;			
Qy	21	YLSAFYAIEFIFGLGNVTVVFGYLCMKNNWS--SNVYFLNLSISDFAFCTLPILIKSY	79
Db	42	FLPLYSILVFLFGLGNSVVVL-VLFYKRLSRMTDVLNLAISDLFLVLSLPFW-GYY	99
Qy	80	ANDKGTGVDVLCISNRVYVLTNLTSTILFTFISMDRYLLMKYPFPEHFLQKKEFALLIS	139
Db	100	AADQWVFLGVLGCKWLSIYLVGFSIGIFFVILMSIDRYLAIVHGVSFMRVRYTYGVITS	159
Qy	140	LAVNALVTLEVLPLMTLINSVPKEEGSNCIDYASSGNPEHNLVYSLCLTLGLFPLISVM	199
Db	160	LATWAVAVFASLPGLLFSTCYTERTNHTSCKTRYSANSTWKLSSLEINILGLVPLIGM	219
Qy	200	CFYKVMVFLKRRSQOQATPLDKPQRLVVLAVVIFSLFTPHIMNLRIRASRLDSW	259
Db	220	-LFCYSMTII---RTLQCKSKKKNKAVKMFVAVVVLFGFWTPYINVLFTYLVELEVL	274
Qy	260	PQGTOKAIKSIYTLTRPLAFNLAINPIFYFLMGDHYREMLISKFR	306
Db	275	QDCSLEKYLDFALQATETLAFIHCLNLPITYFFLGEKFRKIVQLFK	321
RESULT 9			
ID	Q9ERK9	PRELIMINARY;	PRT; 328 AA.
AC	Q9ERK9;		
DT	01-MAR-2001 (TReMBLrel. 16, Created)		
DT	01-MAR-2001 (TReMBLrel. 16, Last sequence update)		
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)		
DE	P2Y6 receptor (Hypothetical 36.7 kDa protein).		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RF*	SEQUENCE FROM N.A.		
RC	STRAIN=129/SVEV;		
RX	MEDLINE=21160052; PubMed=11259526;		
RX	Lazarowski E.R., Rochelle L.G., O'Neal W.K., Ribeiro C.M.P.,		
RA	Grubb B.R., Zhang V., Harden T.K., Boucher R.C.;		

RT "Cloning and functional characterization of two murine uridine
 RT nucleotide receptors reveal a potential target for correcting ion
 RT transport deficiency in cystic fibrosis gallbladder."
 RL J. Pharmacol. Exp. Ther. 297:43-49(2001).
 RN [2]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF298899; AAG24619.1; -

DR EMBL; BC027331; AAG27331.1; -

DR HSP; P34996; IDDP

DR InterPro: IP000276; GPCR_Rhodpsn.

DR Pfam: PF00001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.

KW Hypothetical protein; Receptor.

SQ SEQUENCE 328 AA; 36721 MW; 00F9DF5ADADF903E CRC64;

Query Match 20.5%; Score 338; DB 11; Length 328;
 Best Local Similarity 29.6%; Pred. No. 2.9e-22;
 Matches 84; Conservative 54; Mismatches 136; Indels 10; Gaps 4;

QY 22 LSFAFAIEFIFGLGNVTGVFGYLCMKNNSSNVYLFNLSISDFALCTLPILIKSYA- 80

DB 29 LTPVSVLVGLPUNICVIAICASRRTLTRESVYTLNALADLMYACSLPLIYNAR 88

QY 81 NDKGTGYDVLICISNRYVLHTNLTYSILFTFISMDRYLLMKYPFRE-HFLQKKEFAILIS 139

DB 89 GDHWPFGLACRFVRFANLHGSILFTICISFORYIGICHPLASWHKRGRRAAWVVC 148

QY 140 LAVWALVTLEVLPMLTFTFINSVPKEGSGNCIDYASSGNPEHNLISLCLTLLGFLPLISVM 199

DB 149 GYVWLAVTAQCLPTAVFAATGIGTQRNRTCYDLSPPILSTRYLPYGMALTVIGFLPFLP 208

QY 200 CFYIKVNVFLKRRSQQAATLPL-----DKPQRLVWLVAVVIFSLFTPYHIMRLTAS 254

DB 209 LACYCRMA---RLRCRQDGPAGVQAQRSSKAARMAVVAVFSLFPFHTTAYLAV 265

QY 255 RLDSWPQGTOKAISKIYTLRPLAFLNSAINPIFYFLMGDHYR 298

DB 266 RSTPGVSCPVLTEFAAYKGTGPPFASVNSVLDPLILFYFTQOKFR 309

RESULT 10

Q9MYJ8 PRELIMINARY; PRT; 355 AA.

AC Q9MYJ8

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE Chemokine receptor.

GN CCRI.

OS Callithrix jacchus (Common marmoset).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.

OX NCBI_TaxID=9483;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20153429; PubMed=10686294;

RA Liang M., Rosser M.P., Ng H.P., May K., Bauman J.G., Islam I.,

RA Ghannam A., Kretschmer P.J., Pu H., Dunning L., Snider R.M.,

RA Morrissey M.M., Hesselgesser J., Perez H.D., Horuk R.;

RT "Species selectivity of a small molecule antagonist for the CCR1

chemokine receptor."

RL Eur. J. Pharmacol. 389:41-49(2000).

DR EMBL; AF127528; AAF36453.1; -

DR InterPro: IP000276; GPCR_Rhodpsn.

DR Pfam: PF00001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

KW Receptor.

SQ SEQUENCE 355 AA; 40928 MW; 2B01C47E9874A2C1 CRC64;

Query Match 20.3%; Score 335; DB 6; Length 355;
 Best Local Similarity 28.1%; Pred. No. 5.8e-22;
 Matches 85; Conservative 63; Mismatches 138; Indels 16; Gaps 7;

QY 12 ATEAILNKYLSAFYAIEFIFGLGNVTGVFGYLCMKNNSSNVYLFNLSISDFALCT 71

DB 27 ANERAFGAKLLPPLYSLSVFIGLVGNILVVVVOYKRLKNMTSIYILNLALISDLLEFT 86

QY 72 LPILIKSYANDKGTGYDVLIC--ISNRYVLHTNLTYSILFTFISMDRYLLMKYPFREHFL 129

DB 87 LPFWISYQLKTDWVFGNAMKVLSSFY--YTGLYSEIFFILLITIDRYLAIVHAVPALRA 144

QY 130 QKKKEFAILISLAVWALVTLEVLPMLTFTFINSVPKEGSGNC-IDYASSGNPEHNLISLCLT 188

DB 145 RTVTFGVITISIIWLAIALSLPGLYFAKTQWEITHRTCSLHFPHESSRQEWKLFQALKLN 204

QY 189 LLGLFIPLSVMCFYIKVNVFLKRRSQQAATLPLDKPQRLVWLVAVVIFSLFTPYHIMR 248

DB 205 LLGLVPLVLMVICYTGIIKILLRRPNEKS-----KAVRLFIVIMIFFLFTWTPYNTT 259

QY 249 NLRIASRLDSWPQGTOKAISKIYTLRPLAFLNSAINPIFYFLMGDHYREMLISKFR 307

DB 260 LISVFQDF-LFTYCEQGRQLDLAIQVETMAYTHCCVNPVIYAFYGERFKHL----RQ 314

QY 308 YF 309

DB 315 LF 316

RESULT 11

Q9JLY8 PRELIMINARY; PRT; 355 AA.

AC Q9JLY8

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE Macrophage inflammatory protein-1 alpha receptor.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CD;

RX MEDLINE=20555330; PubMed=11091494;

RA Waller A., Nayee P., Czaplowski L.G.;

RT "Identification and characterization of a rat macrophage inflammatory

protein-1 alpha receptor."

RL J. Hematother. Stem Cell Res. 9:703-710(2000).

DR EMBL; AF119381; AAF34340.1; -

DR InterPro: IP000276; GPCR_Rhodpsn.

DR Pfam: PF00001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.

KW Receptor.

SQ SEQUENCE 355 AA; 40838 MW; 2FEB8661D1E6E075 CRC64;

Query Match 20.2%; Score 333; DB 11; Length 355;

Best Local Similarity 28.6%; Pred. No. 8.7e-22;

Matches 85; Conservative 69; Mismatches 119; Indels 24; Gaps 9;

QY 22 LSFAFAIEFIFGLGNVTGVFGYLCMKNNSSNVYLFNLSISDFALCTLPILIKSYAN 81

DB 37 LPLYSFVFIIGVGNILVILMOHRRQSMTSIYLFNLAVSDLVFLFTLFWIDYKIK 96

QY 82 DKGTGYDVLIC--ISNRYVLHTNLTYSILFTFISMDRYLLMKYPFREHFLQKKEFAILIS 139

DB 97 DNWVFGDAMCKLLSGFYLL--GLYSEIFFILLITIDRYLAIVHAVFSLRARVITGILTS 154

QY 140 LAVWALVTLEVLPMLTFTFINSVPKEGSGNCIDYASSGNPEHNL-----IYSLCLTLGLFLI 194

```
Db 155 IINWALIASIPALCFKQWETHHTC-----SPHPDESCLKTKRQALKNLGLIL 210
Qy 195 PLSVCMFFYKMW-VFLKRRSQOQATALPLDKPQRLVAVVIPSILFTPYHIMNRURIA 253
Db 211 PLLVMICVAGIIRILLRPNRKAQV-----RLFAITLFLLELLTWTYNL--TFVFS 262
Qy 254 SRLD-SWFGQCTQ-KAISKIYTLTPFLAFLNSAINPIFYFLMGDHYREMLISKPROX 308
Db 263 AFQDVLFTNQCEQSKOLDLAIOQTEVIAIYTHCCVNPVIYVGERFKYLRQLFQRH 319

RESULT 12
Q9MYJ9 PRELIMINARY; PRT; 351 AA.
AC Q9MYJ9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Chemokine receptor.
GN CCRL.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20153429; PubMed=10686294;
RA Liang M., Rosser M.P., Ng H.P., May K., Bauman J.G., Islam I.,
RA Ghannam A., Kretschmer P.J., Pu H., Dunning L., Snider R.M.,
RA Morrissey M.M., Hesselgesser J., Perez H.D., Horuk R.;
RT "Species selectivity of a small molecule antagonist for the CCRL
RT chemokine receptor.";
RL Eur. J. Pharmacol. 389:41-49(2000).
DR EMBL: AF127527; AAF36452.1; -.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR_F1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECEPTOR_F1_2; 1.
KW Receptor.
SQ SEQUENCE 351 AA; 40534 MW; B3FED6117A141552 CRC64;

Query Match 19.8%; Score 326; DB 6; Length 351;
Best Local Similarity 27.8%; Pred. No. 3.6e-21;
Matches 83; Conservative 65; Mismatches 137; Indels 14; Gaps 7;

Qy 22 LSFAFYAIEIFGLGNVTGVVFGVLCMKNNNSNVYLFNLISDFAFCTLPILIKSVAN 81
Db 37 LPPLYSLVFVIGLVGNVVLVLMYKRLRSMTSIYLLNLALISDLLFLFTLPFWIDYRLK 96
Qy 82 DKGTGVDVLC--ISNRYVLTNLTYSILFTFTSMRYMLKYPFRHFLOKKEFAILLIS 139
Db 97 DDWVGDVLCVLSGLY--YVGLYSEVFFILLTDIYLAIVHAVFALRARTVSGIVTS 154
Qy 140 LAVWALVTLEVLPMLTFTINSVPKEGNC-IDYASSGNPEHNLITVSLCTLTGLFLPLSV 198
Db 155 IVTWALTILAAIPGRFSKTOWEFTHTCTSLHPHESLRQWQFOALKNLILGLVPLLV 214
Qy 199 MCFYKMWVFLKRRSQOQATALPLDKPQRLVAVVIPSILFTPYHIMNRRIASRLDS 258
Db 215 MVVYCTGTGIIQLLRNNEKKS-----RAVLFIIVMLIFFLFTWTPYNL--TLVSAFQDS 267
Qy 259 -WPGQCTQ-KAISKIYTLTPFLAFLNSAINPIFYFLMGDHYREMLISKPROYFKSITSF 315
Db 268 LFTNQCEQSKOLDLAIOQTEVIAIYTHCCVNPVIYVGERFQYLRQLFHTYLAQWLPF 326

RESULT 13
Q912H4 PRELIMINARY; PRT; 360 AA.
ID Q912H4
AC Q912H4;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
```

```
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE C-C chemokine receptor 4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEW;
RA Garcia G.E., Chen S., Xia Y., Harrison J., Wilson C.B., Johnson R.J.,
RA Bacon K.B., Feng L.;
RT "Mononuclear cell-infiltrate inhibition by blocking macrophage-derived
RT chemokine results in attenuation of developing crescentic
RT glomerulonephritis.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF432872; AAL30398.1; -.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR01559; DUFFYANTIGEN.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR_F1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECEPTOR_F1_2; 1.
KW Receptor.
SQ SEQUENCE 360 AA; 41218 MW; 5095C6CD299EIF8B CRC64;

Query Match 19.4%; Score 320; DB 11; Length 360;
Best Local Similarity 29.4%; Pred. No. 1.3e-20;
Matches 88; Conservative 58; Mismatches 133; Indels 20; Gaps 9;

Qy 21 YLSAFYAIEIFGLGNVTGVVFGVLCMKNNNS-SNVYLFNLISDFAFCTLPILIKSY 79
Db 41 FLPLYSLVFVGLGNGSVVVL-VLFYKRLKSWTDVYLLNLALISDLLFLVSLPFW-GY 98
Qy 80 ANDKGTGVDVLCISNRYVLTNLTYSILFTFTSMRYMLKYPFRHFLOKKEFAILLIS 139
Db 99 AADQWVEGLGLCKIISWMLVGVFGYSGIFFILMSIDRYLAIVHAVFSLRARTLTGYVTS 158
Qy 140 LAVWALVTLEVLPMLTFTINSVPKEGNCIDYASSGNPEHNLITVSLCTLTGLFLPLSV 199
Db 159 LITWSAVFASPLGLLSTCDTENHNYCKYQYSVNSTTKWVLSLEINVLGLVPLGIM 218
Qy 200 CFYKMWVFLKRRSQOQATALPLDKPQRLVAVVIPSILFTPYHIMNRRIASRLDSW 259
Db 219 -LFCYSMLITFLRHCKNEKK---NRAVRMTFAVVVFLGFWTPYNNVFLVLELVEVL 273
Qy 260 PQGCT-QKAISKIYTLTPFLAFLNSAINPIFYFLMGDHYREMLISKPROYFKSITSPT 317
Db 274 -QDCTERYLDYATQATETATTAIFHCCLPNPVIYFLGE-----KPKYIAQL--PRT 321

RESULT 14
Q8R528 PRELIMINARY; PRT; 309 AA.
ID Q8R528
AC Q8R528;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Cysteinyl leukotriene 2 receptor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishii S.;
RT "Mouse CysLT2 Gene.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB058930; BAB86881.1; -.
KW Receptor.
SQ SEQUENCE 309 AA; 35226 MW; 96FACC6B8AF96974 CRC64;

Query Match 19.3%; Score 318.5; DB 11; Length 309;
Best Local Similarity 27.3%; Pred. No. 1.5e-20;
Matches 84; Conservative 56; Mismatches 147; Indels 21; Gaps 8;
```

```

QY 3 QNLSCENLATEAILNKYLSAFYAIEFIFGLLGNTVVFVGYLFCKMKNSSNVYLFNLS 62
Db 13 RNCIENF-----KKEFPYIYLIIFWALGNCFSYVFLQCKKSTSVNFMNLNA 65

QY 63 ISDFAFICTLPILIKSY-ANDKTYGDVLCISNRYVLTHTNLKTSILFTFISMDRYLLMK 121
Db 66 TSDFLFISTLPFRADYFGRSNNIFGDLACRVMSYSLYVNMVTSYVFLVLSVVRFLATV 125

QY 122 YPEREHFLOKKEFAILISLAWALVTLVPLMLTFINSVPKEGNS---CIDYASSGNPE 178
Db 126 HPRFMFHTSVRSAILWGLIIVFIMAS--SALLLVNG--QEEKDNIIISCLE--LSPQKEK 180

QY 179 HNLISLCLTLLGLFLPSVMCFYFYKMV-VFLKRRSQOQATALPLDKPQRLWLAVVIF 237
Db 181 SLLIMNHIAVAGFLLPELTLITCYILLIIRILLKAIPESGPRAHRKALTIVIAMITF 240

QY 238 SILFTPYHIMRNRIASRLDSNPQGTQKAIKSIYTLTRPLAFLNSAINPIFYFLMGDHY 297
Db 241 LLCFLPYHALRTLHLV---TWDKDCGDLVHKATVITLTMAAANSCFNPFLYYFAGENF 296

QY 298 REMLISKF 305
Db 297 KARLRAIF 304

```

RESULT 15

```

Q95LH1
ID Q95LH1 PRELIMINARY; PRT; 355 AA.
AC Q95LH1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE G protein-coupled receptor 1
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Wade-Evans A.M., Javan C., Russell J., Jenkins A.;
RT "Cloning and sequencing of simian G-protein coupled receptors, which
RT may function as SIV/SHIV co-receptors, from cynomolgus macaque
RT PBMCs.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292382; AAK97052.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 355 AA; 41370 MW; 1DDE41BE9CAC525C CRC64;

```

```

Query Match 19.2%; Score 316.5; DB 6; Length 355;
Best Local Similarity 29.2%; Pred.No. 2.6e-20;
Matches 86; Conservative 59; Mismatches 119; Indels 31; Gaps 9;

```

```

QY 26 YAIIEIFGLGNTVVFVGYLFCKMKNWS--SNVYLFNLSIDFAPLCTLPILIKSYA--N 81
Db 45 YCLSEVLGIPGNAIVW---FTGFKWKRTVSTLWFLNLAIADFIQLFLPLYI-SYVVMN 100

QY 82 DKGTYGDVLCISNRYVLTHTNLKTSILFTFISMDRYLLMKYPPFEHFLOKKEFAILISLA 141
Db 101 FHWPGIWLCKANFTAOQNNMFASVFFLTVISLDHYTHLIHPVLSHRHRTLKNSLIIVIF 160

QY 142 VVALVLEVLPLTFINSVPKEGNSCIDYASSGNPEHNLISLCLT----LIGFLIPLS 197
Db 161 IWLASLIGGPGALYFRTVFNENHTLVCNNFQKHDPDLTVIRHVLVWVYVIGFLPPL 220

QY 198 VMCFYKMKVFLKRRSQOQATALPLDKPQRLVAVVI-FSILFTPYHIMRNRIASRL 256
Db 221 TMSICYLCLILKVKRS-----ILSSRHFWTILAVVAVVVCWTPYHLFSIWELTIH 274

```

```

QY 257 DSMPOGCTQKAIKSIYTLTRPLAFLNSAINPIFYFLMGDHYREMLISKFRQYFKS 311
Db 275 NSYSHHVMOAGI-----PLSTGLAFLNSCLNPILYVLIS-----KKFOARFRS 317

```

Search completed: April 2, 2003, 15:19:21
Job time : 44 secs

